

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 14:00:58 ; Search time 50.2093 Seconds
(without alignments)
213.658 Million cell updates/sec

Title: US-10-010-667A-2_COPY_185_218

Perfect score: 192

Sequence: 1 RRSYRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 34

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mbc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	192	100.0	338	6 Q9GL50	Q9GL50 sus scrofa
2	187	97.4	339	11 Q924J9	Q924J9 mus musculus
3	187	97.4	339	11 Q9CWR7	Q9CWR7 mus musculus
4	139	72.4	339	11 Q924Z2	Q924Z2 mus musculus
5	110	57.3	489	11 Q8BWB6	Q8BWB6 mus musculus
6	109	56.8	488	11 Q99P41	Q99P41 rattus norv
7	109	56.8	488	11 Q8CI59	Q8CI59 mus musculus
8	109	56.8	514	11 Q924Z1	Q924Z1 mus musculus
9	109	56.8	526	11 Q80ZP3	Q80ZP3 mus musculus
10	107	55.7	488	4 Q9NVB5	Q9NVB5 homo sapien
11	107	55.7	488	4 Q86Sf6	Q86Sf6 homo sapien
12	107	55.7	490	4 Q8NFT2	Q8NFT2 homo sapien
13	107	55.7	490	4 Q8IUE7	Q8IUE7 homo sapien
14	107	55.7	498	4 Q7Z389	Q7Z389 homo sapien
15	105	54.7	488	11 Q8C5F0	Q8C5F0 mus musculus
16	96.5	50.3	456	4 Q8TDP3	Q8TDP3 homo sapien

17	96.5	50.3	487	4 Q8TF03	Q8TF03 homo sapien
18	96.5	50.3	487	4 Q8NEW6	Q8NEW6 homo sapien
19	62.5	32.6	480	16 Q67595	Q67595 aquifex aeo
20	62	32.3	366	16 Q8A6D0	Q8A6D0 bacteroides
21	60	31.2	926	4 Q8TE49	Q8TE49 homo sapien
22	60	31.2	926	11 Q8R554	Q8R554 mus musculus
23	58.5	30.5	1027	5 Q95QC7	Q95QC7 caenorhabdi
24	58	30.2	213	16 Q813H4	Q813H4 bacillus ce
25	57.5	29.9	556	11 Q8CFS0	Q8CFS0 mus musculus
26	57.5	29.9	858	4 Q9NQ53	Q9NQ53 homo sapien
27	56	29.2	150	12 Q9YVJ4	Q9YVJ4 melanoplus
28	56	29.2	230	2 Q9RFP3	Q9RFP3 mycoplasma
29	55.5	28.9	502	2 Q48904	Q48904 microcystis
30	55	28.6	81	16 Q8IMC0	Q8IMC0 bacillus an
31	55	28.6	1379	5 Q817P4	Q817P4 dictyosteli
32	54	28.1	81	16 Q812U5	Q812U5 bacillus ce
33	53.5	27.9	427	4 Q8WWA7	Q8WWA7 homo sapien
34	53.5	27.9	831	16 Q8D8X5	Q8D8X5 vibrio vuln
35	53	27.6	152	17 Q30299	Q30299 archaeoglob
36	53	27.6	307	16 Q8EAP3	Q8EAP3 shewanella
37	53	27.6	336	2 Q87519	Q87519 escherichia
38	53	27.6	391	16 Q8FEB8	Q8FEB8 escherichia
39	53	27.6	429	17 Q8TFE0	Q8TFE0 methanosarc
40	53	27.6	536	16 Q89Q05	Q89Q05 bradyrhizob
41	52.5	27.3	533	16 Q82MD6	Q82MD6 streptomyce
42	52.5	27.3	685	4 Q81WK5	Q81WK5 homo sapien
43	52	27.1	132	15 Q91WQ0	Q91WQ0 human immun
44	52	27.1	252	4 Q8N2A8	Q8N2A8 homo sapien
45	52	27.1	252	4 Q8N5Y1	Q8N5Y1 homo sapien

ALIGNMENTS

RESULT 1

Q9GL50
ID Q9GL50 PRELIMINARY; PRT; 338 AA.
AC Q9GL50; (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Six transmembrane endothelial antigen of PAEC.
OS Sus scrofa (Pig).
OC Rukaryota; Eutazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Theria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Nagasaka T., Boulday G., Coupel S., Coulon F., Tesson L.,
RA Heslan J.-M., Soullou J.-P., Chaireau B.;
RT "Differential gene expression in endothelial cells during TNF-alpha-
RT and LPS-mediated activation."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF319659; AAG33868.1; --
DR GO: GO:0016021; C:integral to membrane; IEA.
KW Transmembrane.
SQ SEQUENCE 338 AA; 39918 MW; ED490E86E067A32B CRC64;

Query Match 100.0%; Score 192; DB 6; Length 338;
Best Local Similarity 100.0%; Pred. No. 7.9e-17;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSYRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 34

Db 184 RRSYRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 217

RESULT 2

Q924J9
ID Q924J9 PRELIMINARY; PRT; 339 AA.
AC Q924J9
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

```

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Six-transmembrane epithelial antigen of the prostate.
GN STEAP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
EN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryonic stem cells;
RX MEDLINE=21371909; PubMed=11479226;
RA Yang D., Holt G.E., Velders M.P., Kwon E.D., Kast W.M.;
RT "Murine six-transmembrane epithelial antigen of the prostate, prostate
RT stem cell antigen, and prostate-specific membrane antigen: prostate-
RT specific cell-surface antigens highly expressed in prostate cancer of
RT transgenic adenocarcinoma mouse prostate mice.";
RL Cancer Res. 61:5857-5860(2001).
DR EMBL; AF297098; AAK83126.1; -.
DR MGD; MGI:1917608; Steap.
DR GO; GO:0016021; C:integral to membrane; IEA.
KW Transmembrane.
SQ SEQUENCE 339 AA; 39318 MW; 4B26A71FF59E84F CRC64;

Query Match          97.4%; Score 187; DB 11; Length 339;
Best Local Similarity 94.1%; Pred. No. 3.5e-16;
Matches 32; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRSRYKLLNWAYQVQVQNKEDAWIEHDVWRMEI 34
   |||||:|||||:|||||:|||||:|||||:
Db 185 RRSRYKLLNWAYQVQVQNKEDAWIEHDVWRMEI 218

RESULT 3
Q9CWR7 ID Q9CWR7 PRELIMINARY; PRT; 339 AA.
AC Q9CWR7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE 241007B19Rik protein.
GN STEAP OR 241007B19RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
EN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryonic stem cells;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itch M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Straubli F., Suzuki R., Tonita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wysshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK010437; BAB26938.1; -.
DR MGD; MGI:1917608; Steap.
SQ SEQUENCE 339 AA; 39264 MW; 3F7AB9C7520F0968 CRC64;
```

```

Query Match          97.4%; Score 187; DB 11; Length 339;
Best Local Similarity 94.1%; Pred. No. 3.5e-16;
Matches 32; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRSRYKLLNWAYQVQVQNKEDAWIEHDVWRMEI 34
   |||||:|||||:|||||:|||||:|||||:
Db 185 RRSRYKLLNWAYQVQVQNKEDAWIEHDVWRMEI 218

RESULT 4
Q924Z2 ID Q924Z2 PRELIMINARY; PRT; 339 AA.
AC Q924Z2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dudulin.
GN STEAP OR 1010001D01RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
EN [1]
RP SEQUENCE FROM N.A.
RA Seru V., Manivet P., Lamblin D., Vaubourdolle M., Kellermann O.,
RA Loric S.;
RT "Prostate and non-prostate expression of dudulin, the mouse ortholog
RT of human STEAP.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY029584; AAK50537.1; -.
DR MGD; MGI:1917608; Steap.
SQ SEQUENCE 339 AA; 39109 MW; 32A2C29F2E33BD0 CRC64;

Query Match          72.4%; Score 139; DB 11; Length 339;
Best Local Similarity 73.5%; Pred. No. 5.8e-10;
Matches 25; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 RRSRYKLLNWAYQVQVQNKEDAWIEHDVWRMEI 34
   |||||:|||||:|||||:|||||:|||||:
Db 185 RRSRYDLNLAVKQVLANKEDAWIEHDVWRMEI 218

RESULT 5
Q8BWB6 ID Q8BWB6 PRELIMINARY; PRT; 489 AA.
AC Q8BWB6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Weakly similar to tumor suppressor PHYE.
GN 4921538B17RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
EN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK052981; BAC35230.1; -.
DR MGD; MGI:1921301; 4921538B17RIK.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR004455; NADPoxred F420.
DR Pfam; PF03807; F420_oxidored; 1.
SQ SEQUENCE 489 AA; 55760 MW; 98CD63D59DDDF24C CRC64;

Query Match          57.3%; Score 110; DB 11; Length 489;
Best Local Similarity 61.8%; Pred. No. 4.8e-06;
```

[illegible]

0:

Search completed: July 12, 2004, 14:10:25
Job time : 51.2093 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 13:52:02 ; Search time 10.6744 Seconds
(without alignments)

165.853 Million cell updates/sec

Title: US-10-010-667a-2_COPY_185_218

Perfect score: 192

Sequence: 1 RRSRYKLINWAYQQQNKEDAWIEHDVWRMEI 34

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	192	100.0	339	1 STEA_HUMAN	Q9u8e8 homo sapien
2	58.5	30.5	1418	1 STEA_CAEEL	P34641 caenorhabdi
3	58	30.2	900	1 GGAB_BACSU	P46918 bacillus su
4	54.5	28.4	636	1 SYT_WIGER	Q8d3c0 wigglewort
5	54	28.1	630	1 Y4BJ_RHISN	P55377 rhizobium s
6	52	27.1	79	1 YOKC_BACSU	P54566 bacillus su
7	52	27.1	1813	1 UN13_CAEEL	P27715 caenorhabdi
8	51	26.6	495	1 GLPK_SINY3	P74260 synecocyst
9	51	26.6	1017	1 ACAD_ARATH	Q911k7 arabidopsis
10	50.5	26.3	313	1 TRUEB_PROMA	P59880 prochloroco
11	50	26.0	784	1 ALP4_SCHPO	Q9y705 schizosacch
12	50	26.0	1574	1 SYJ1_RAT	Q62910 rattus norv
13	50	26.0	5430	1 MACF_HUMAN	Q9upn3 homo sapien
14	50	26.0	5938	1 MAC4_HUMAN	Q9rn28 bacillus an
15	49.5	25.8	95	1 YC04_BACAN	P57323 buchnera ap
16	49.5	25.8	274	1 DAPD_BUCAI	Q58669 methanococ
17	49.5	25.8	350	1 YC73_METJA	Q84253 chlamydia t
18	49.5	25.8	787	1 OXAA_CHLTR	Q92987 chlamydia p
19	49.5	25.8	940	1 SYV_CHLPN	Q9pdk6 xylella fas
20	49	25.5	257	1 TRUA_XYLPA	Q87d51 xylella fas
21	49	25.5	257	1 TRUA_XYLFT	Q87d51 xylella fas
22	49	25.5	500	1 GLPK_ANASP	Q8y05 anabaena sp
23	49	25.5	512	1 Y4WA_RHISN	P55679 rhizobium s
24	49	25.5	837	1 NCM2_HUMAN	O15394 homo sapien
25	49	25.5	837	1 NCM2_MOUSE	O35136 mus musculu
26	49	25.5	1324	1 SYJ1_BOVIN	O18964 bos taurus
27	49	25.5	1575	1 SYJ1_HUMAN	O43426 homo sapien
28	48.5	25.3	266	1 IF2A_SULSO	Q97z79 sulfobolus
29	48.5	25.3	339	1 MO2L_CAEEL	Q9tzm2 caenorhabdi
30	48	25.0	286	1 YAY8_SCHPO	O10216 schizosacch
31	48	25.0	405	1 ASSY_PSEPK	P59604 pseudomonas
32	48	25.0	466	1 SYN_SHEON	Q8eez1 shewanella
33	48	25.0	620	1 SYR_BIFLO	Q8g4v2 bifidobacte

RESULT 1

ID	STEA_HUMAN	STANDARD;	PRT;	339 AA.
AC	Q9UHE8; O95034;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Six transmembrane epithelial antigen of prostate.			
GN	STEAP OR STEAP1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20056277; PubMed=10588738;			
RA	Hubert R.S., Vivanco I., Chen E., Rastegar S., Leong K., Mitchell S.C., Madraswala R., Zhou Y., Kuo J., Raitano A.B., Jakobovits A., Safran D.C., Afar D.E.H.;			
RT	"STEAP: a prostate-specific cell-surface antigen highly expressed in human prostate tumors.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 96:14523-14528(1999).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Abu-Threideh J., Stoneking T., Langston Y., Maupin R.;			
RL	Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Skin;			
RX	MEDLINE=22398257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Udgin T.B., Toshiyuki S., Carninci P., Prange C., Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Marra M.A.;			
RA	Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
CC	- SUBCELLULAR LOCATION: Integral membrane protein (Potential).			
CC	- TISSUE SPECIFICITY: Highly expressed in prostate tumors.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its			

Q92442 mucor javan
P31627 caprine arc
Q87dp4 vibrio para
O06994 bacillus su
P15977 escherichia
P27175 gluconobact
O02740 bos taurus
O61460 caenorhabdi
Q9abm9 caulobacter
Q48332 halobacteri
P59825 rhodopirell
Q92331 saccharomyc

ALIGNMENTS

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AF186249; AAF17479.1; -;
DR EMBL; AC005053; AAC79150.1; ALT_INIT.
DR EMBL; AC004969; AAD15620.2; -;
DR EMBL; BC011802; AAH11802.1; -;
DR Genew; HGNC:11378; STEAP.
DR MIM; 604415; -;
DR GO; GO:0005887; C:integral to plasma membrane; TAS.
DR GO; GO:0005911; C:intercellular junction; TAS.
DR GO; GO:0015267; F:channel/pore class transporter activity; TAS.
DR KW Transmembrane; Antigen.
FT TRANSMEM 71 91 POTENTIAL.
FT TRANSMEM 119 139 POTENTIAL.
FT TRANSMEM 164 184 POTENTIAL.
FT TRANSMEM 218 238 POTENTIAL.
FT TRANSMEM 258 278 POTENTIAL.
FT TRANSMEM 291 311 POTENTIAL.
SQ SEQUENCE 339 AA; 55443A170C870387 CRC64;

Query Match 100.0%; Score 192; DB 1; Length 339;
Best Local Similarity 100.0%; Pred. No. 5.8e-19;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSYRYKLLNWAYQVOQNKEDAWIEHDVWRMEI 34
|||||
DB 185 RRSYRYKLLNWAYQVOQNKEDAWIEHDVWRMEI 218

RESULT 2

CELL1_CABEL STANDARD; PRT; 1418 AA.
ID CELL1_CABEL
AC P34641;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ced-11 protein.
DE Ced-11 OR ZK512.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=Bristol N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Connell M., Copsey T., Cooper J., Coulson A.,
RA Bonfield J., Burton J., Durbin R., Favello A., Fraser A.,
RA Craxton M., Dear S., Du Z., Durbin R., Hawkins T., Hillier L., Jier M.,
RA Fulton L., Gardner A., Green P., Jenkins T., Kesteven N.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laissner N.,
RA Lacroix P., Lighning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
RA Sims M., Smaldon N., Smith A., Smith M., Sonnenhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstein L., Wilkinson-Sproat J.,
RA Wohlman P.,
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RL Nature 368:32-38(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z22177; CAA80145.1; -;

DR PIR; S40764; S40764.
DR WormPep; ZK512.3; CE00409.
DR InterPro; IPR002111; Cat_channel_TrpL.
SQ SEQUENCE 1418 AA; 159237 MW; 4FC83B9E7ADF7788 CRC64;

Query Match 30.5%; Score 58.5; DB 1; Length 1418;
Best Local Similarity 43.5%; Pred. No. 5.5;
Matches 10; Conservative 5; Mismatches 5; Indels 3; Gaps 1;

QY 15 QVOQNKED---AWIEHDVWRMEI 34
||:|:| |||||:
DB 1355 EVAENNDTDNAWIEHDVWJSL 1377

RESULT 3

GGAB_BACSU STANDARD; PRT; 900 AA.
ID GGAB_BACSU
AC P46918;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Minor teichoic acid biosynthesis protein ggaB.
GN GGAB OR BSU35680.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RC SEQUENCE FROM N.A.
RP Freymond P.-P., Karamata D.;
RC STRAIN=168;
RA "Sequencing and analysis of two gga genes associated with the
RT synthesis of the minor teichoic acid of Bacillus subtilis 168."
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RC SEQUENCE FROM N.A.
RP STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Etian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hulio M.P., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaert-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapius A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Taconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256(1997).
CC -I- FUNCTION: Involved in the biosynthesis of galactosamine-containing
CC minor teichoic acid, a non-essential cell wall polymer in
CC B. subtilis 168.
CC -I- PATHWAY: Poly(glucopyranosyl N-acetyl-galactosamine 1-phosphate)
CC teichoic acid biosynthesis.
CC

InterPro: IPR006195; tRNA_Ligase II.
DR Pfam: PF00129; HGTP-anticodon; 1.
DR Pfam: PF00587; tRNA-synt 2b; 1.
DR PRINTS: PR01047; TRNASYNTHTR.
DR TIGRFAMS: TIGR00418; thrS; 1.
DR PROSITE: PS50862; AA tRNA_Ligase II; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Metal-binding; Zinc; Complete proteome.
FT DOMAIN 242 533 CATALYTIC.
FT METAL 333 333 ZINC (CATALYTIC) (BY SIMILARITY) .
FT METAL 384 384 ZINC (CATALYTIC) (BY SIMILARITY) .
FT METAL 510 510 ZINC (CATALYTIC) (BY SIMILARITY) .
SQ SEQUENCE 636 AA; 75381 MW; 585695AB61D57P7E CRC64;

Query Match 28.4%; Score 54.5; DB 1; Length 636;
Best Local Similarity 42.9%; Pred.No. 7.9;
Matches 12; Conservative 5; Mismatches 8; Indels 3; Gaps 1;

QY 5 RYKLLNWAYQQVQONKEDAWIEHDVWRM 32
| | | : | | | : | | | : | | | :
DB 281 RKXKKFYQEV---KTPCIIDHDLWRL 305

RESULT 5
Y4BU_RHISN STANDARD; PRT; 630 AA.
ID Y4BU_RHISN AC P53377;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical 67.9 kDa protein Y4BU.
GN Y4BU.
OS Rhizobium sp. (strain NGR234) .
OG Plasmid sym pNGR234a.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX NCBI_Taxid=394;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=97305956; PubMed=9163424;
RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
RA Perret X.;
RT "Molecular basis of symbiosis between Rhizobium and legumes.";
RL Nature 387:394-401(1997).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: DISTANTLY RELATED TO PEPTIDASE FAMILY S2C.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB000666; AAB91625.1; -.
CC MEROPS; S01.UPC; -.
DR InterPro: IPR009003; Cys Ser trypsin.
DR InterPro: IPR001254; Peptidase_S1.
DR InterPro: IPR001940; Peptidase_SIC.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS; PR00834; PROTEASES2C.
KW Hypothetical protein; plasmid; Transmembrane.
FT TRANSMEM 8 28 POTENTIAL.
FT TRANSMEM 258 278 POTENTIAL.
SQ SEQUENCE 630 AA; 67854 MW; 4C1FB7BA2EB44A02 CRC64;

Query Match 28.1%; Score 54; DB 1; Length 630;
Best Local Similarity 35.5%; Pred.No. 9.1;
Matches 11; Conservative 5; Mismatches 7; Indels 8; Gaps 1;

QY 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVWR 31
| | | : | | | : | | | : | | | :
| | | : | | | : | | | : | | | :

Dd 296 RRYRRELFGWEYR-----AKLEPBIWR 318

RESULT 6

YQKC_BACSU

ID YQKC_BACSU STANDARD; PRT; 79 AA.

AC P54566;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Hypothetical protein yqkC.

GN YQKC OR BSU23650.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=1423;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=168 / JH642;

RX MEDLINE=97124195; PubMed=8969508;

RA Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M.,

RA Kobayashi Y.

RT "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of

RT the Bacillus subtilis genome containing the skin element and many

RT sporulation genes."

RL Microbiology 142:3103-3111(1996).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RX MEDLINE=98044033; PubMed=9384377;

RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,

RA Azavedo V., Bertoletti M.G., Bessieres P., Bolotin A., Borchert S.,

RA Borris R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,

RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,

RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,

RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,

RA Enric K.D., Errington J., Fabret C., Ferrari E., Foulger D.,

RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,

RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,

RA Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,

RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,

RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,

RA Kobayashi Y., Koetter P., Konigstein G., Krohn S., Kumano M.,

RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,

RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,

RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,

RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,

RA Parro V., Pohl T.M., Portetelle D., Portwollik S., Prescott A.M.,

RA Presecan E., Fujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,

RA Rieger M., Rivoita C., Roche E., Roche B., Rose M., Sadale Y.,

RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,

RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,

RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,

RA Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A.,

RA Tosato V., Uchiyama S., Vandenbol M., Vannier P., Vassarotti A.,

RA Viari A., Wambut R., Wedler E., Wedler H., Weitzenger T.,

RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,

RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,

RT "The complete genome sequence of the Gram-positive bacterium Bacillus

RT subtilis."

RL Nature 390:249-256(1997).

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; D84432; BAA12635.1; --

CC EMBL; Z99116; CAB14297.1; --

CC PIR; E69966; E69966.

CC Subtilist; E611758; yqkC.

CC

KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 79 AA; 9287 MW; 34852BDAEA4805D2 CRC64;

Query Match 27.1%; Score 52; DB 1; Length 79;
Best Local Similarity 28.1%; Pred. NO. 1.7;
Matches 9; Conservative 8; Mismatches 15; Indels 0; Gaps 0;

QY 2 RSYRYKLNWAYQQVQVKEDAWLEHVDWRME 33
Db 23 QNHPTLLHWSIGGAESIKKDVLLQDEMTPF 54

RESULT 7

UN13 CAEEL

ID UN13 CAEEL STANDARD; PRT; 1813 AA.

AC P27715; Q17665; Q23512; Q81095;

DT 01-AUG-1992 (Rel. 23, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Phorbol ester/diacylglycerol-binding protein unc-13 (Uncoordinated

DE protein 13)

GN UNC-13 OR ZK524.2.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORM 2).

RC STRAIN=Bristol N2;

RX MEDLINE=91288538; PubMed=2062851;

RA Maruyama I.N., Brenner S.

RT "A phorbol ester/diacylglycerol-binding protein encoded by the unc-13

RT gene of Caenorhabditis elegans."

Proc. Natl. Acad. Sci. U.S.A. 88:5729-5733(1991).

[2]

RN PARTIAL SEQUENCE FROM N.A. (ISOFORMS 1 AND 3).

RC STRAIN=Bristol N2;

RX MEDLINE=20483775; PubMed=11029047;

RA Kohn R.E., Duerr J.S., McManus J.R., Duke A., Rakow T.L., Maruyama H.,

RA Moulder G., Maruyama I.N., Barstead R.J., Rand J.B.

RT "Expression of multiple UNC-13 proteins in the Caenorhabditis elegans

RT nervous system."

Mol. Biol. Cell 11:3441-3452(2000).

[3]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RA Gardner A.E., Lloyd C.R.

RT Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.

[4]

RP REVISIONS, AND ALTERNATIVE SPLICING.

RA Durbin R.

RP Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.

[5]

RN ZINC AND PHORBOL-ESTERS BINDING.

RP MEDLINE=93075060; PubMed=1445255;

RA Ahmed S., Maruyama I.N., Kozma R., Lee J., Brenner S., Lim L.

RT "The Caenorhabditis elegans unc-13 gene product is a phospholipid-

RT dependent high-affinity phorbol ester receptor."

Biochem. J. 287:995-999(1992).

CC -!- FUNCTION: May form part of a signal transduction pathway,

CC transducing the signal from diacylglycerol to effector functions.

CC One such function could be the release of neurotransmitter from

CC neurons.

CC -!- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=5;

CC Name=1; Synonyms=ZK524.2a;

CC IsoId=P27715-1; Sequence=Displayed;

CC Name=2;

CC IsoId=P27715-2; Sequence=VSP_004484;

CC Name=3;

CC IsoId=P27715-3; Sequence=VSP_004485;

CC Name=4; Synonyms=ZK524.2b;

CC IsoId=P27715-4; Sequence=VSP_004486, VSP_004487;

CC

[illegible]

```

CC
EMBL; D90913; BAA18354.1; -.
DR PUR; S75895; S75895.
DR HSPF; P08659; IGLC.
DR HAWAP; MF 00186; -.
DR HAWAP; MF 00186; -.
DR InterPro; IPR000577; FGGY_kin.
DR InterPro; IPR005999; Glycerol_kin.
DR Pfam; PF00370; FGGY; 1.
DR Pfam; PF02782; FGGY_C; 1.
DR TIGRfams; TIGR01311; glycerol_kin; 1.
DR PROSITE; PS00933; FGGY_KINASES_1; FALSE_NEG.
DR PROSITE; PS00445; FGGY_KINASES_2; 1.
DR KW Glycerol metabolism; Transferase; Kinase; ATP-binding;
DR Complete proteome.
DR NP_BIND 156 168 ATP (PROBABLE).
DR NP_BIND 495 AA; 54351 MW; 71AAC71E76816736 CRC64;
DR SEQUENCE
Query Match 26.6%; Score 51; DB 1; Length 495;
Best Local Similarity 34.4%; Pred.No. 18;
Matches 11; Conservative 7; Mismatches 10; Indels 4; Gaps 2;

QY 4 YRYKLLNWAYQQVQONKEDA-WIEHD---VWR 31
      :: ||::: | |::: | |::: | |
Db 26 YEGNIVGQAYKEITQFYFKAGWVEHDALEIWR 57

RESULT 9
ACAD ARATH
ID ACAD_ARATH STANDARD; PRT; 1017 AA.
AC QLLK7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Potential calcium-transporting ATPase 13, plasma membrane-type
DE (EC 3.6.3.8) (Ca(2+)-ATPase isoform 13).
GN ACA13 OR AT3G22910 OR F5N5.8.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsi.
ON NCBI_TaxID=3702;
RX [1]
RC SEQUENCE FROM N.A.
RP STRAIN=cv. Columbia;
RX MEDLINE=20363099; PubMed=10907853;
RA Kaneko T., Katoh T., Sato S., Nakamura A., Asamizu E., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3 II.
RT Sequence features of the 4,251,695 bp regions covered by 90 Pl, TAC
RT and BAC clones.";
RL DNA Res. 7:217-221(2000).
CC -! FUNCTION: This magnesium-dependent enzyme catalyzes the hydrolysis
CC of ATP coupled with the translocation of calcium from the cytosol
CC out of the cell or into organelles (By similarity).
CC -! CATALYTIC ACTIVITY: ATP + H(2)O + Ca(2+) (Cis) = ADP + phosphate +
CC Ca(2+) (Trans).
CC -! ENZYME REGULATION: Activated by calmodulin (By similarity).
CC -! SUBCELLULAR LOCATION: Integral membrane protein.
CC -! DOMAIN: The N-terminus contains an autoinhibitory calmodulin-
CC binding domain, which binds calmodulin in a calcium-dependent
CC fashion (By similarity).
CC -! SIMILARITY: Belongs to the cation transport ATPases family (P-type
CC ATPases). Subfamily IIB.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC

```



```

RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritsch C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Fumelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880 (2002).
CC -1- FUNCTION: Component of the gamma tubule complex that is required
CC for the regulation of both interphase microtubules and mitotic
CC bipolar spindles.
CC -1- SUBCELLULAR LOCATION: Spindle pole body and the microtubule
CC organizing center (MTOC).
CC -1- SIMILARITY: Belongs to the GCP family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL078627; CAB44767.1; -.
CC EMBL; AB026664; BAA77269.1; -.
CC FJ; T43510; T43510.
CC DR GeneDB-SPombe; SPBC365.15; -.
CC DR InterPro; IPR007259; SPC97_Spc98.
CC DR Pfam; PF04130; SPC97_Spc98; 1.
CC KW Microtubule; Mitosis.
CC SEQUENCE 784 AA; 90157 MW; E529CE217FFDA2B2 CRC64;
CC -----
Query Match 26.0%; Score 50; DB 1; Length 784;
Best Local Similarity 31.2%; Pred.No. 41;
Matches 10; Conservative 8; Mismatches 8; Indels 6; Gaps 2;
QY 4 YRYKLINWAYQQYQONKEDAWIEH--DVWRM 32
Db 592 FRFLL---LRHVMQLQNSWVQHSKNSAWRL 620
RESULT 12
SYJL RAT
ID _SYJL RAT STANDARD; PRT; 1574 AA.
AC Q62910; Q62910; Q62911; Q61028;
DT 30-MAY-2000 (Rel. 39, Created)
DR 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Synaptotagmin 1 (EC 3.1.3.36) (Synaptic inositol-1,4,5-trisphosphate 5-
DE phosphatase 1).
GN SYNJ1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RP STRAIN=Sprague-Dawley; TISSUE=Brain;
RC MEDLINE=96149250; PubMed=8552192;
RA McPherson P.S., Garcia E.P., Slepnev V.I., David C., Zhang X.,
RA Grabs D., Sosin W.S., Bauerfeind R., Nemoto Y., De Camilli P.;
RT "A presynaptic inositol-5-phosphatase."
RL Nature 379:353-357 (1996).

```

RP SEQUENCE FROM N.A. (ISOFORM 3).
RC TISSUE=Brain;
RX MEDLINE=98374013; PubMed=9710239;
RA Woscholski R., Finan P.M., Radley E., Parker P.J.;
RT "Identification and characterisation of a novel splice variant of
RT synaptotagmin.";
RL FEBS Lett. 432:5-8(1998).
RN [3]
RP ALTERNATIVE SPLICING.
RC TISSUE=Brain;
RX MEDLINE=96394655; PubMed=9798761;
RA Ranjaun A.R., McPherson P.S.;
RT "Tissue-specific alternative splicing generates two synaptotagmin
RT isoforms with differential membrane binding properties.";
RL J. Biol. Chem. 271:24856-24861(1996).
RN [4]
RP INTERACTION WITH SH3GL1, SH3GL2 AND SH3GL3.
RC TISSUE=Brain;
RX MEDLINE=97385143; PubMed=9238017;
RA Ringsstad N., Nemoto Y., De Camilli P.;
RT "The SH3p4/SH3p8/SH3p3 protein family: binding partners for
RT synaptotagmin and dynamin via a Grb2-like Src homology 3 domain.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:8569-8574(1997).
RN [5]
RP -!- FUNCTION: Inositol 5-phosphatase which has a role in clathrin-
CC mediated endocytosis.
CC -!- CATALYTIC ACTIVITY: 1-phosphatidyl-1D-myo-inositol 4,5-
CC bisphosphate + H(2)O = 1-phosphatidyl-1D-myo-inositol 4-phosphate
CC + phosphate.
CC -!- SUBUNIT: Binds AMPH, SH3GL1, SH3GL2 and SH3GL3.
CC -!- SUBCELLULAR LOCATION: Localized mainly in the cytoplasm (By
CC similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=6;
CC Name=1; Synonyms=170 kDa;
CC Note=A stop codon in position 1309 is suppressed due to
CC alternative splicing;
CC Name=2; Synonyms=145 kDa;
CC IsoId=Q62910-2; Sequence=VSP_002686;
CC Name=3; Synonyms=Delta-SAC;
CC IsoId=Q62910-3; Sequence=VSP_002684;
CC Name=4; Synonyms=170 kDa-16AA; Sequence=VSP_002685;
CC IsoId=Q62910-4; Sequence=VSP_002685;
CC Name=5; Synonyms=145 kDa-16AA;
CC IsoId=Q62910-5; Sequence=VSP_002685; VSP_002686;
CC Name=6; Synonyms=Delta-SAC-16AA;
CC IsoId=Q62910-6; Sequence=VSP_002684, VSP_002685;
CC -!- TISSUE SPECIFICITY: Isoform 1 is found in neonatal brain, and in a
CC wide variety of adult non-neuronal tissues. Isoform 2 is expressed
CC predominantly in the neurons, but is also found in all other
CC tissues at much lower levels. Isoform 1 and isoform 2 are detected
CC in the lung and heart. Isoform 1 is expressed at higher levels
CC than isoform 2 in the testis and liver and both isoforms are not
CC detected in the skeletal muscle. Isoform 3 with the 16-amino-acid
CC insert is only found in the brain while isoform 3 without the 16-
CC amino-acid insert is found in the lung.
CC -!- DEVELOPMENTAL STAGE: At embryonic day 12 (E12) only isoform 1 is
CC seen while at E16 and E18 isoform 1 and isoform 2 are seen. In the
CC adult brain expression of isoform 2 increases dramatically as
CC compared with its expression in embryonic brain where as isoform 1
CC decreases to undetectable levels.
CC -!- DOMAIN: Binds to EPS15 (a clathrin coat-associated protein) via a
CC C-terminal domain containing three Asn-Pro-Phe (NPP) repeats (By
CC similarity).
CC -!- DOMAIN: The C-terminal proline-rich region mediates binding to a
CC variety of SH3 domain-containing proteins including AMPH, SH3GL1,
CC SH3GL2, SH3GL3 and GRB2.
CC -!- DOMAIN: Splicing of the SAC1 domain does not alter the catalytic
CC activity of synaptotagmin 1.
CC -!- SIMILARITY: In the central section; belongs to the inositol-1,4,5-
CC triphosphate 5-phosphatase family.
CC -!- SIMILARITY: Contains 1 SAC domain.

CC -!- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U45479; AAB60525.1; -.
DR EMBL; U91836; AAC24807.1; -.
DR EMBL; AJ006855; CAA07267.1; ALT_TERM.
DR InterPro; IPR005135; Exo_endo_phos.
DR InterPro; IPR000300; IPPC.
DR InterPro; IPR000504; RNA_rec_mot.
DR InterPro; IPR002013; SYJA_N.
DR Pfam; PF03372; Exo_endo_phos; 1.
DR Pfam; PF02383; SYJA_N; 1.
DR SMART; SMO0128; IPPC; 1.
DR PROSITE; PS00102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
DR PROSITE; PS00275; SAC; 1.
DR Hydrolase; Alternative splicing; Repeat; Endocytosis; RNA-binding;
KW Multigene family.
FT DOMAIN 119 442 SAC
FT DOMAIN 500 899 CATALYTIC (POTENTIAL).
FT DOMAIN 894 971 RNA-BINDING (RRM).
FT DOMAIN 900 1574 PRO-RICH.
FT DOMAIN 1033 1036 POLY-SER.
FT DOMAIN 1123 1110 POLY-PRO.
FT DOMAIN 1125 1126 POLY-PRO.
FT DOMAIN 1537 1545 POLY-PRO.
FT DOMAIN 1401 1423 3 X 3 AA REPEATS OF N-P-F.
FT REPEAT 1401 1403 1.
FT REPEAT 1410 1412 2.
FT REPEAT 1421 1423 3.
FT VARSPLIC 1 400 Missing (in isoform 3 and isoform 6).
FT VARSPLIC 1140 1155 /FTId=VSP_002684.
FT VARSPLIC 1309 1574 Missing (in isoform 4, isoform 5 and isoform 6).
FT VARSPLIC 1574 /FTId=VSP_002685.
FT CONFLICT 588 588 Missing (in isoform 2 and isoform 5).
FT CONFLICT 1140 1140 G -> D (IN REF. 1; AAB60525).
FT CONFLICT 1156 1156 G -> GG (IN REF. 2).
FT CONFLICT 1574 1574 MISSING (IN REF. 2).
SQ SEQUENCE 1574 AA; 172880 MW; 9DEESC876EAB3ADF CRC64;
Query Match 26.0%; Score 50; DB 1; Length 1574;
Best Local Similarity 27.6%; Pred. No. 90;
Matches 8; Conservative 8; Mismatches 13; Indels 0; Gaps 0;
QY 3 SYRYKLNWAYQQVQONKEDATIEHDVWR 31
Db 783 TYKYDLFSDYDTSEKCRTPAWTDRLWR 811
RESULT 13
MACF_HUMAN
ID MACF_HUMAN STANDARD; PRT; 5430 AA.
AC Q9UPN3; O75053; Q8WXY2; Q9H540; Q9UKP0; Q9ULG9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Microtubule-actin crosslinking factor 1, isoforms 1/2/3 (Actin cross-
DE linking family protein 7) (Macrophilin 1) (Trabeculin-alpha) (620 kDa
DE actin-binding protein) (ABP620).
GN MACF1 OR AC7 OR ABP620 OR KIAA0465 OR KIAA1251.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;


```
FT DOMAIN 3418 3482 COILED COIL (POTENTIAL) .
FT DOMAIN 3596 3666 COILED COIL (POTENTIAL) .
FT DOMAIN 3786 3806 COILED COIL (POTENTIAL) .
FT DOMAIN 3852 3931 COILED COIL (POTENTIAL) .
FT DOMAIN 3967 3997 COILED COIL (POTENTIAL) .
FT DOMAIN 4084 4218 COILED COIL (POTENTIAL) .
FT DOMAIN 4343 4378 COILED COIL (POTENTIAL) .
FT DOMAIN 4408 4437 COILED COIL (POTENTIAL) .
FT DOMAIN 4468 4498 COILED COIL (POTENTIAL) .
FT DOMAIN 4907 4935 COILED COIL (POTENTIAL) .
FT DOMAIN 5044 5067 COILED COIL (POTENTIAL) .
FT REPEAT 314 355 SPECTRIN 1.
FT REPEAT 591 623 SPECTRIN 2.
FT REPEAT 680 784 SPECTRIN 3.
FT REPEAT 786 800 SPECTRIN 4.
FT DOMAIN 871 923 SH3.
FT REPEAT 1250 1272 SPECTRIN 5.
FT REPEAT 1287 1342 SPECTRIN 6.
FT REPEAT 1455 1534 SPECTRIN 7.
FT REPEAT 1547 1659 SPECTRIN 8.
FT REPEAT 1815 1891 SPECTRIN 9.
FT REPEAT 1932 2042 SPECTRIN 10.
FT REPEAT 2260 2280 SPECTRIN 11.
FT REPEAT 2372 2395 SPECTRIN 12.
FT REPEAT 2398 2507 SPECTRIN 13.
FT REPEAT 2510 2618 SPECTRIN 14.
FT REPEAT 2621 2728 SPECTRIN 15.
FT REPEAT 2731 2838 SPECTRIN 16.
FT REPEAT 2841 2945 SPECTRIN 17.
FT REPEAT 2987 3024 SPECTRIN 18.
FT REPEAT 3136 3163 SPECTRIN 19.
FT REPEAT 3187 3274 SPECTRIN 20.
FT REPEAT 3277 3383 SPECTRIN 21.
FT REPEAT 3386 3492 SPECTRIN 22.
FT REPEAT 3495 3601 SPECTRIN 23.
FT REPEAT 3604 3673 SPECTRIN 24.
FT REPEAT 3713 3819 SPECTRIN 25.
FT REPEAT 3832 3927 SPECTRIN 26.
FT REPEAT 3982 4043 SPECTRIN 27.
FT REPEAT 4046 4152 SPECTRIN 28.
FT REPEAT 4155 4262 SPECTRIN 29.

Query Match 26.0%; Score 50; DB 1; Length 5430;
Best Local Similarity 40.7%; Pred. No. 3.6e+02;
Matches 11; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 5 RYKLNWAYQQVQNKEDAWIEHDVWR 31
DB 5048 FQRLNDALDRLEELKEFANFDVWR 5074

RESULT 14
MAC4 HUMAN
ID MAC4 HUMAN STANDARD; PRT; 5938 AA.
AC Q96PK2; Q9WXY1;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Microtubule-actin crosslinking factor 1, isoform 4.
GN MACF1 OR ACF7 OR ABP620 OR KIAA0465 OR KIAA1251.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
RX MEDLINE=21833812; PubMed=11845288;
RA Gong T.-W.L., Besirli C.G., Lomax M.I.;
RT "MACF1 gene structure: a hybrid of plectin and dystrophin.";
RL Mamm. Genome 12:852-861(2001).
CC -!- FUNCTION: May play a role in cross-linking cytoskeletal proteins
CC by binding intermediate filaments to the N-terminal plectin
CC repeats and microtubules to the C-terminus.
```

```
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Name=4;
CC IsoId=Q96PK2-1; Sequence=Displayed;
CC Name=1;
CC IsoId=Q9UPN3-1; Sequence=External;
CC Name=2;
CC IsoId=Q9UPN3-2; Sequence=External;
CC Name=3;
CC IsoId=Q9UPN3-3; Sequence=External;
CC -!- TISSUE SPECIFICITY: Expressed in heart, lung, pituitary and
CC placenta, not found in brain, kidney, liver, pancreas or skeletal
CC muscle.
CC -!- SIMILARITY: Belongs to the plectin or cytolinker family.
CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC -!- SIMILARITY: Contains 19 plectrin repeats.
CC -!- SIMILARITY: Contains 32 spectrin repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF317696; AAL09459.1; -.
CC EMBL; AF325341; AAL38997.1; -.
CC EMBL; AF325333; AAL38997.1; JOINED.
CC EMBL; AF325334; AAL38997.1; JOINED.
CC EMBL; AF325335; AAL38997.1; JOINED.
CC EMBL; AF325336; AAL38997.1; JOINED.
CC EMBL; AF325339; AAL38997.1; JOINED.
CC EMBL; AF325340; AAL38997.1; JOINED.
CC DR GO; GO:0005856; C:cytoskeleton; ISS.
CC DR GO; GO:0005509; F:calcium ion binding; ISS.
CC DR GO; GO:0008017; F:microtubule binding; ISS.
CC DR InterPro; IPR002048; EF-hand.
CC DR InterPro; IPR003108; GAS2.
CC DR InterPro; IPR001101; Plectrin_repeat.
CC DR InterPro; IPR002017; Spectrin.
CC DR Pfam; PF00036; ehand; 2.
CC DR Pfam; PF02187; GAS2; 1.
CC DR Pfam; PF00681; Plectrin; 11.
CC DR Pfam; PF00435; spectrin; 26.
CC DR ProDom; PDOM00012; EF-hand; 1.
CC DR SMART; SM00054; EfH; 2.
CC DR SMART; SM00243; GAS2; 1.
CC DR SMART; SM00250; PLEC; 19.
CC DR SMART; SM00150; SPEC; 32.
CC DR PROSITE; PS00018; EF_HAND; 2.
KW Cytoskeleton; Calcium; Calcium-binding; Repeat; Coiled coil;
KW Alternative splicing.
KW DOMAIN 1830 1936 COILED COIL (POTENTIAL) .
FT DOMAIN 2001 2192 COILED COIL (POTENTIAL) .
FT DOMAIN 2282 2345 COILED COIL (POTENTIAL) .
FT DOMAIN 2477 2507 COILED COIL (POTENTIAL) .
FT DOMAIN 2541 2654 COILED COIL (POTENTIAL) .
FT DOMAIN 2686 2814 COILED COIL (POTENTIAL) .
FT DOMAIN 2887 2919 COILED COIL (POTENTIAL) .
FT DOMAIN 3046 3197 COILED COIL (POTENTIAL) .
FT DOMAIN 3262 3503 COILED COIL (POTENTIAL) .
FT DOMAIN 3632 3666 COILED COIL (POTENTIAL) .
FT DOMAIN 3746 3779 COILED COIL (POTENTIAL) .
FT DOMAIN 3920 3984 COILED COIL (POTENTIAL) .
FT DOMAIN 4098 4168 COILED COIL (POTENTIAL) .
FT DOMAIN 4288 4308 COILED COIL (POTENTIAL) .
FT DOMAIN 4354 4386 COILED COIL (POTENTIAL) .
FT DOMAIN 4397 4433 COILED COIL (POTENTIAL) .
FT DOMAIN 4469 4489 COILED COIL (POTENTIAL) .
FT DOMAIN 4586 4720 COILED COIL (POTENTIAL) .
FT DOMAIN 4845 4880 COILED COIL (POTENTIAL) .
```

```
FT DOMAIN 4910 4939 COILED COIL (POTENTIAL).
FT DOMAIN 4970 5000 COILED COIL (POTENTIAL).
FT DOMAIN 5409 5437 COILED COIL (POTENTIAL).
FT DOMAIN 5546 5569 COILED COIL (POTENTIAL).
FT CA_BIND 5598 5610 EF-HAND 1 (POTENTIAL).
FT CA_BIND 5634 5646 EF-HAND 2 (POTENTIAL).
FT REPEAT 12 49 PLECTIN 1.
FT REPEAT 53 88 PLECTIN 2.
FT REPEAT 89 126 PLECTIN 3.
FT REPEAT 130 164 PLECTIN 4.
FT REPEAT 166 202 PLECTIN 5.
FT REPEAT 203 240 PLECTIN 6.
FT REPEAT 243 278 PLECTIN 7.
FT REPEAT 279 316 PLECTIN 8.
FT REPEAT 318 354 PLECTIN 9.
FT REPEAT 354 388 PLECTIN 10.
FT REPEAT 400 434 PLECTIN 11.
FT REPEAT 434 468 PLECTIN 12.
FT REPEAT 468 502 PLECTIN 13.
FT REPEAT 502 536 PLECTIN 14.
FT REPEAT 536 570 PLECTIN 15.
FT REPEAT 570 604 PLECTIN 16.
FT REPEAT 604 638 PLECTIN 17.
FT REPEAT 638 672 PLECTIN 18.
FT REPEAT 672 706 PLECTIN 19.
FT REPEAT 706 740 PLECTIN 20.
FT REPEAT 740 774 PLECTIN 21.
FT REPEAT 774 808 PLECTIN 22.
FT REPEAT 808 842 PLECTIN 23.
FT REPEAT 842 876 PLECTIN 24.
FT REPEAT 876 910 PLECTIN 25.
FT REPEAT 910 944 PLECTIN 26.
FT REPEAT 944 978 PLECTIN 27.
FT REPEAT 978 1012 PLECTIN 28.
FT REPEAT 1012 1046 PLECTIN 29.
FT REPEAT 1046 1080 PLECTIN 30.
FT REPEAT 1080 1114 PLECTIN 31.
FT REPEAT 1114 1148 PLECTIN 32.
FT REPEAT 1148 1182 PLECTIN 33.
FT REPEAT 1182 1216 PLECTIN 34.
FT REPEAT 1216 1250 PLECTIN 35.
FT REPEAT 1250 1284 PLECTIN 36.
FT REPEAT 1284 1318 PLECTIN 37.
FT REPEAT 1318 1352 PLECTIN 38.
FT REPEAT 1352 1386 PLECTIN 39.
FT REPEAT 1386 1420 PLECTIN 40.
FT REPEAT 1420 1454 PLECTIN 41.
FT REPEAT 1454 1488 PLECTIN 42.
FT REPEAT 1488 1522 PLECTIN 43.
FT REPEAT 1522 1556 PLECTIN 44.
FT REPEAT 1556 1590 PLECTIN 45.
FT REPEAT 1590 1624 PLECTIN 46.
FT REPEAT 1624 1658 PLECTIN 47.
FT REPEAT 1658 1692 PLECTIN 48.
FT REPEAT 1692 1726 PLECTIN 49.
FT REPEAT 1726 1760 PLECTIN 50.
FT REPEAT 1760 1794 PLECTIN 51.
FT REPEAT 1794 1828 PLECTIN 52.
FT REPEAT 1828 1862 PLECTIN 53.
FT REPEAT 1862 1896 PLECTIN 54.
FT REPEAT 1896 1930 PLECTIN 55.
FT REPEAT 1930 1964 PLECTIN 56.
FT REPEAT 1964 1998 PLECTIN 57.
FT REPEAT 1998 2032 PLECTIN 58.
FT REPEAT 2032 2066 PLECTIN 59.
FT REPEAT 2066 2100 PLECTIN 60.
FT REPEAT 2100 2134 PLECTIN 61.
FT REPEAT 2134 2168 PLECTIN 62.
FT REPEAT 2168 2202 PLECTIN 63.
FT REPEAT 2202 2236 PLECTIN 64.
FT REPEAT 2236 2270 PLECTIN 65.
FT REPEAT 2270 2304 PLECTIN 66.
FT REPEAT 2304 2338 PLECTIN 67.
FT REPEAT 2338 2372 PLECTIN 68.
FT REPEAT 2372 2406 PLECTIN 69.
FT REPEAT 2406 2440 PLECTIN 70.
FT REPEAT 2440 2474 PLECTIN 71.
FT REPEAT 2474 2508 PLECTIN 72.
FT REPEAT 2508 2542 PLECTIN 73.
FT REPEAT 2542 2576 PLECTIN 74.
FT REPEAT 2576 2610 PLECTIN 75.
FT REPEAT 2610 2644 PLECTIN 76.
FT REPEAT 2644 2678 PLECTIN 77.
FT REPEAT 2678 2712 PLECTIN 78.
FT REPEAT 2712 2746 PLECTIN 79.
FT REPEAT 2746 2780 PLECTIN 80.
FT REPEAT 2780 2814 PLECTIN 81.
FT REPEAT 2814 2848 PLECTIN 82.
FT REPEAT 2848 2882 PLECTIN 83.
FT REPEAT 2882 2916 PLECTIN 84.
FT REPEAT 2916 2950 PLECTIN 85.
FT REPEAT 2950 2984 PLECTIN 86.
FT REPEAT 2984 3018 PLECTIN 87.
FT REPEAT 3018 3052 PLECTIN 88.
FT REPEAT 3052 3086 PLECTIN 89.
FT REPEAT 3086 3120 PLECTIN 90.
FT REPEAT 3120 3154 PLECTIN 91.
FT REPEAT 3154 3188 PLECTIN 92.
FT REPEAT 3188 3222 PLECTIN 93.
FT REPEAT 3222 3256 PLECTIN 94.
FT REPEAT 3256 3290 PLECTIN 95.
FT REPEAT 3290 3324 PLECTIN 96.
FT REPEAT 3324 3358 PLECTIN 97.
FT REPEAT 3358 3392 PLECTIN 98.
FT REPEAT 3392 3426 PLECTIN 99.
FT REPEAT 3426 3460 PLECTIN 100.
FT REPEAT 3460 3494 PLECTIN 101.
FT REPEAT 3494 3528 PLECTIN 102.
FT REPEAT 3528 3562 PLECTIN 103.
FT REPEAT 3562 3596 PLECTIN 104.
FT REPEAT 3596 3630 PLECTIN 105.
FT REPEAT 3630 3664 PLECTIN 106.
FT REPEAT 3664 3698 PLECTIN 107.
FT REPEAT 3698 3732 PLECTIN 108.
FT REPEAT 3732 3766 PLECTIN 109.
FT REPEAT 3766 3800 PLECTIN 110.
FT REPEAT 3800 3834 PLECTIN 111.
FT REPEAT 3834 3868 PLECTIN 112.
FT REPEAT 3868 3902 PLECTIN 113.
FT REPEAT 3902 3936 PLECTIN 114.
FT REPEAT 3936 3970 PLECTIN 115.
FT REPEAT 3970 4004 PLECTIN 116.
FT REPEAT 4004 4038 PLECTIN 117.
FT REPEAT 4038 4072 PLECTIN 118.
FT REPEAT 4072 4106 PLECTIN 119.
FT REPEAT 4106 4140 PLECTIN 120.
FT REPEAT 4140 4174 PLECTIN 121.
FT REPEAT 4174 4208 PLECTIN 122.
FT REPEAT 4208 4242 PLECTIN 123.
FT REPEAT 4242 4276 PLECTIN 124.
FT REPEAT 4276 4310 PLECTIN 125.
FT REPEAT 4310 4344 PLECTIN 126.
FT REPEAT 4344 4378 PLECTIN 127.
FT REPEAT 4378 4412 PLECTIN 128.
FT REPEAT 4412 4446 PLECTIN 129.
FT REPEAT 4446 4480 PLECTIN 130.
FT REPEAT 4480 4514 PLECTIN 131.
FT REPEAT 4514 4548 PLECTIN 132.
FT REPEAT 4548 4582 PLECTIN 133.
FT REPEAT 4582 4616 PLECTIN 134.
FT REPEAT 4616 4650 PLECTIN 135.
FT REPEAT 4650 4684 PLECTIN 136.
FT REPEAT 4684 4718 PLECTIN 137.
FT REPEAT 4718 4752 PLECTIN 138.
FT REPEAT 4752 4786 PLECTIN 139.
FT REPEAT 4786 4820 PLECTIN 140.
FT REPEAT 4820 4854 PLECTIN 141.
FT REPEAT 4854 4888 PLECTIN 142.
FT REPEAT 4888 4922 PLECTIN 143.
FT REPEAT 4922 4956 PLECTIN 144.
FT REPEAT 4956 4990 PLECTIN 145.
FT REPEAT 4990 5024 PLECTIN 146.
FT REPEAT 5024 5058 PLECTIN 147.
FT REPEAT 5058 5092 PLECTIN 148.
FT REPEAT 5092 5126 PLECTIN 149.
FT REPEAT 5126 5160 PLECTIN 150.
FT REPEAT 5160 5194 PLECTIN 151.
FT REPEAT 5194 5228 PLECTIN 152.
FT REPEAT 5228 5262 PLECTIN 153.
FT REPEAT 5262 5296 PLECTIN 154.
FT REPEAT 5296 5330 PLECTIN 155.
FT REPEAT 5330 5364 PLECTIN 156.
FT REPEAT 5364 5398 PLECTIN 157.
FT REPEAT 5398 5432 PLECTIN 158.
FT REPEAT 5432 5466 PLECTIN 159.
FT REPEAT 5466 5500 PLECTIN 160.
FT REPEAT 5500 5534 PLECTIN 161.
FT REPEAT 5534 5568 PLECTIN 162.
FT REPEAT 5568 5602 PLECTIN 163.
FT REPEAT 5602 5636 PLECTIN 164.
FT REPEAT 5636 5670 PLECTIN 165.
FT REPEAT 5670 5704 PLECTIN 166.
FT REPEAT 5704 5738 PLECTIN 167.
FT REPEAT 5738 5772 PLECTIN 168.
FT REPEAT 5772 5806 PLECTIN 169.
FT REPEAT 5806 5840 PLECTIN 170.
FT REPEAT 5840 5874 PLECTIN 171.
FT REPEAT 5874 5908 PLECTIN 172.
FT REPEAT 5908 5942 PLECTIN 173.
FT REPEAT 5942 5976 PLECTIN 174.
FT REPEAT 5976 6010 PLECTIN 175.
FT REPEAT 6010 6044 PLECTIN 176.
FT REPEAT 6044 6078 PLECTIN 177.
FT REPEAT 6078 6112 PLECTIN 178.
FT REPEAT 6112 6146 PLECTIN 179.
FT REPEAT 6146 6180 PLECTIN 180.
FT REPEAT 6180 6214 PLECTIN 181.
FT REPEAT 6214 6248 PLECTIN 182.
FT REPEAT 6248 6282 PLECTIN 183.
FT REPEAT 6282 6316 PLECTIN 184.
FT REPEAT 6316 6350 PLECTIN 185.
FT REPEAT 6350 6384 PLECTIN 186.
FT REPEAT 6384 6418 PLECTIN 187.
FT REPEAT 6418 6452 PLECTIN 188.
FT REPEAT 6452 6486 PLECTIN 189.
FT REPEAT 6486 6520 PLECTIN 190.
FT REPEAT 6520 6554 PLECTIN 191.
FT REPEAT 6554 6588 PLECTIN 192.
FT REPEAT 6588 6622 PLECTIN 193.
FT REPEAT 6622 6656 PLECTIN 194.
FT REPEAT 6656 6690 PLECTIN 195.
FT REPEAT 6690 6724 PLECTIN 196.
FT REPEAT 6724 6758 PLECTIN 197.
FT REPEAT 6758 6792 PLECTIN 198.
FT REPEAT 6792 6826 PLECTIN 199.
FT REPEAT 6826 6860 PLECTIN 200.
FT REPEAT 6860 6894 PLECTIN 201.
FT REPEAT 6894 6928 PLECTIN 202.
FT REPEAT 6928 6962 PLECTIN 203.
FT REPEAT 6962 6996 PLECTIN 204.
FT REPEAT 6996 7030 PLECTIN 205.
FT REPEAT 7030 7064 PLECTIN 206.
FT REPEAT 7064 7098 PLECTIN 207.
FT REPEAT 7098 7132 PLECTIN 208.
FT REPEAT 7132 7166 PLECTIN 209.
FT REPEAT 7166 7200 PLECTIN 210.
FT REPEAT 7200 7234 PLECTIN 211.
FT REPEAT 7234 7268 PLECTIN 212.
FT REPEAT 7268 7302 PLECTIN 213.
FT REPEAT 7302 7336 PLECTIN 214.
FT REPEAT 7336 7370 PLECTIN 215.
FT REPEAT 7370 7404 PLECTIN 216.
FT REPEAT 7404 7438 PLECTIN 217.
FT REPEAT 7438 7472 PLECTIN 218.
FT REPEAT 7472 7506 PLECTIN 219.
FT REPEAT 7506 7540 PLECTIN 220.
FT REPEAT 7540 7574 PLECTIN 221.
FT REPEAT 7574 7608 PLECTIN 222.
FT REPEAT 7608 7642 PLECTIN 223.
FT REPEAT 7642 7676 PLECTIN 224.
FT REPEAT 7676 7710 PLECTIN 225.
FT REPEAT 7710 7744 PLECTIN 226.
FT REPEAT 7744 7778 PLECTIN 227.
FT REPEAT 7778 7812 PLECTIN 228.
FT REPEAT 7812 7846 PLECTIN 229.
FT REPEAT 7846 7880 PLECTIN 230.
FT REPEAT 7880 7914 PLECTIN 231.
FT REPEAT 7914 7948 PLECTIN 232.
FT REPEAT 7948 7982 PLECTIN 233.
FT REPEAT 7982 8016 PLECTIN 234.
FT REPEAT 8016 8050 PLECTIN 235.
FT REPEAT 8050 8084 PLECTIN 236.
FT REPEAT 8084 8118 PLECTIN 237.
FT REPEAT 8118 8152 PLECTIN 238.
FT REPEAT 8152 8186 PLECTIN 239.
FT REPEAT 8186 8220 PLECTIN 240.
FT REPEAT 8220 8254 PLECTIN 241.
FT REPEAT 8254 8288 PLECTIN 242.
FT REPEAT 8288 8322 PLECTIN 243.
FT REPEAT 8322 8356 PLECTIN 244.
FT REPEAT 8356 8390 PLECTIN 245.
FT REPEAT 8390 8424 PLECTIN 246.
FT REPEAT 8424 8458 PLECTIN 247.
FT REPEAT 8458 8492 PLECTIN 248.
FT REPEAT 8492 8526 PLECTIN 249.
FT REPEAT 8526 8560 PLECTIN 250.
FT REPEAT 8560 8594 PLECTIN 251.
FT REPEAT 8594 8628 PLECTIN 252.
FT REPEAT 8628 8662 PLECTIN 253.
FT REPEAT 8662 8696 PLECTIN 254.
FT REPEAT 8696 8730 PLECTIN 255.
FT REPEAT 8730 8764 PLECTIN 256.
FT REPEAT 8764 8798 PLECTIN 257.
FT REPEAT 8798 8832 PLECTIN 258.
FT REPEAT 8832 8866 PLECTIN 259.
FT REPEAT 8866 8900 PLECTIN 260.
FT REPEAT 8900 8934 PLECTIN 261.
FT REPEAT 8934 8968 PLECTIN 262.
FT REPEAT 8968 9002 PLECTIN 263.
FT REPEAT 9002 9036 PLECTIN 264.
FT REPEAT 9036 9070 PLECTIN 265.
FT REPEAT 9070 9104 PLECTIN 266.
FT REPEAT 9104 9138 PLECTIN 267.
FT REPEAT 9138 9172 PLECTIN 268.
FT REPEAT 9172 9206 PLECTIN 269.
FT REPEAT 9206 9240 PLECTIN 270.
FT REPEAT 9240 9274 PLECTIN 271.
FT REPEAT 9274 9308 PLECTIN 272.
FT REPEAT 9308 9342 PLECTIN 273.
FT REPEAT 9342 9376 PLECTIN 274.
FT REPEAT 9376 9410 PLECTIN 275.
FT REPEAT 9410 9444 PLECTIN 276.
FT REPEAT 9444 9478 PLECTIN 277.
FT REPEAT 9478 9512 PLECTIN 278.
FT REPEAT 9512 9546 PLECTIN 279.
FT REPEAT 9546 9580 PLECTIN 280.
FT REPEAT 9580 9614 PLECTIN 281.
FT REPEAT 9614 9648 PLECTIN 282.
FT REPEAT 9648 9682 PLECTIN 283.
FT REPEAT 9682 9716 PLECTIN 284.
FT REPEAT 9716 9750 PLECTIN 285.
FT REPEAT 9750 9784 PLECTIN 286.
FT REPEAT 9784 9818 PLECTIN 287.
FT REPEAT 9818 9852 PLECTIN 288.
FT REPEAT 9852 9886 PLECTIN 289.
FT REPEAT 9886 9920 PLECTIN 290.
FT REPEAT 9920 9954 PLECTIN 291.
FT REPEAT 9954 9988 PLECTIN 292.
FT REPEAT 9988 10022 PLECTIN 293.
FT REPEAT 10022 10056 PLECTIN 294.
FT REPEAT 10056 10090 PLECTIN 295.
FT REPEAT 10090 10124 PLECTIN 296.
FT REPEAT 10124 10158 PLECTIN 297.
FT REPEAT 10158 10192 PLECTIN 298.
FT REPEAT 10192 10226 PLECTIN 299.
FT REPEAT 10226 10260 PLECTIN 300.
FT REPEAT 10260 10294 PLECTIN 301.
FT REPEAT 10294 10328 PLECTIN 302.
FT REPEAT 10328 10362 PLECTIN 303.
FT REPEAT 10362 10396 PLECTIN 304.
FT REPEAT 10396 10430 PLECTIN 305.
FT REPEAT 10430 10464 PLECTIN 306.
FT REPEAT 10464 10498 PLECTIN 307.
FT REPEAT 10498 10532 PLECTIN 308.
FT REPEAT 10532 10566 PLECTIN 309.
FT REPEAT 10566 10600 PLECTIN 310.
FT REPEAT 10600 10634 PLECTIN 311.
FT REPEAT 10634 10668 PLECTIN 312.
FT REPEAT 10668 10702 PLECTIN 313.
FT REPEAT 10702 10736 PLECTIN 314.
FT REPEAT 10736 10770 PLECTIN 315.
FT REPEAT 10770 10804 PLECTIN 316.
FT REPEAT 10804 10838 PLECTIN 317.
FT REPEAT 10838 10872 PLECTIN 318.
FT REPEAT 10872 10906 PLECTIN 319.
FT REPEAT 10906 10940 PLECTIN 320.
FT REPEAT 10940 10974 PLECTIN 321.
FT REPEAT 10974 11008 PLECTIN 322.
FT REPEAT 11008 11042 PLECTIN 323.
FT REPEAT 11042 11076 PLECTIN 324.
FT REPEAT 11076 11110 PLECTIN 325.
FT REPEAT 11110 11144 PLECTIN 326.
FT REPEAT 11144 11178 PLECTIN 327.
FT REPEAT 11178 11212 PLECTIN 328.
FT REPEAT 11212 11246 PLECTIN 329.
FT REPEAT 11246 11280 PLECTIN 330.
FT REPEAT 11280 11314 PLECTIN 331.
FT REPEAT 11314 11348 PLECTIN 332.
FT REPEAT 11348 11382 PLECTIN 333.
FT REPEAT 11382 11416 PLECTIN 334.
FT REPEAT 11416 11450 PLECTIN 335.
FT REPEAT 11450 11484 PLECTIN 336.
FT REPEAT 11484 11518 PLECTIN 337.
FT REPEAT 11518 11552 PLECTIN 338.
FT REPEAT 11552 11586 PLECTIN 339.
FT REPEAT 11586 11620 PLECTIN 340.
FT REPEAT 11620 11654 PLECTIN 341.
FT REPEAT 11654 11688 PLECTIN 342.
FT REPEAT 11688 11722 PLECTIN 343.
FT REPEAT 11722 11756 PLECTIN 344.
FT REPEAT 11756 11790 PLECTIN 345.
FT REPEAT 11790 11824 PLECTIN 346.
FT REPEAT 11824 11858 PLECTIN 347.
FT REPEAT 11858 11892 PLECTIN 348.
FT REPEAT 11892 11926 PLECTIN 349.
FT REPEAT 11926 11960 PLECTIN 350.
FT REPEAT 11960 11994 PLECTIN 351.
FT REPEAT 11994 12028 PLECTIN 352.
FT REPEAT 12028 12062 PLECTIN 353.
FT REPEAT 12062 12096 PLECTIN 354.
FT REPEAT 12096 12130 PLECTIN 355.
FT REPEAT 12130 12164 PLECTIN 356.
FT REPEAT 12164 12198 PLECTIN 357.
FT REPEAT 12198 12232 PLECTIN 358.
FT REPEAT 12232 12266 PLECTIN 359.
FT REPEAT 12266 12300 PLECTIN 360.
FT REPEAT 12300 12334 PLECTIN 361.
FT REPEAT 12334 12368 PLECTIN 362.
FT REPEAT 12368 12402 PLECTIN 363.
FT REPEAT 12402 12436 PLECTIN 364.
FT REPEAT 12436 12470 PLECTIN 365.
FT REPEAT 12470 12504 PLECTIN 366.
FT REPEAT 12504 12538 PLECTIN 367.
FT REPEAT 12538 12572 PLECTIN 368.
FT REPEAT 12572 12606 PLECTIN 369.
FT REPEAT 12606 12640 PLECTIN 370.
FT REPEAT 12640 12674 PLECTIN 371.
FT REPEAT 12674 12708 PLECTIN 372.
FT REPEAT 12708 12742 PLECTIN 373.
FT REPEAT 12742 12776 PLECTIN 374.
FT REPEAT 12776 12810 PLECTIN 375.
FT REPEAT 12810 12844 PLECTIN 376.
FT REPEAT 12844 12878 PLECTIN 377.
FT REPEAT 12878 12912 PLECTIN 378.
FT REPEAT 12912 12946 PLECTIN 379.
FT REPEAT 12946 12980 PLECTIN 380.
FT REPEAT 12980 13014 PLECTIN 381.
FT REPEAT 13014 13048 PLECTIN 382.
FT REPEAT 13048 13082 PLECTIN 383.
FT REPEAT 13082 13116 PLECTIN 384.
FT REPEAT 13116 13150 PLECTIN 385.
FT REPEAT 13150 13184 PLECTIN 386.
FT REPEAT 13184 13218 PLECTIN 387.
FT REPEAT 13218 13252 PLECTIN 388.
FT REPEAT 13252 13286 PLECTIN 389.
FT REPEAT 13286 13320 PLECTIN 390.
FT REPEAT 13320 13354 PLECTIN 391.
FT REPEAT 13354 13388 PLECTIN 392.
FT REPEAT 13388 13422 PLECTIN 393.
FT REPEAT 13422 13456 PLECTIN 394.
FT REPEAT 13456 13490 PLECTIN 395.
FT REPEAT 13490 13524 PLECTIN 396.
FT REPEAT 13524 13558 PLECTIN 397.
FT REPEAT 13558 13592 PLECTIN 398.
FT REPEAT 13592 13626 PLECTIN 399.
FT REPEAT 13626 13660 PLECTIN 400.
FT REPEAT 13660 13694 PLECTIN 401.
FT REPEAT 13694 13728 PLECTIN 402.
FT REPEAT 13728 13762 PLECTIN 403.
FT REPEAT 13762 13796 PLECTIN 404.
FT REPEAT 13796 13830 PLECTIN 405.
FT REPEAT 13830 13864 PLECTIN 406.
FT REPEAT 13864 13898 PLECTIN 407.
FT REPEAT 13898 13932 PLECTIN 408.
FT REPEAT 13932 13966 PLECTIN 409.
FT REPEAT 13966 14000 PLECTIN 410.
FT REPEAT 14000 14034 PLECTIN 411.
FT REPEAT 14034 14068 PLECTIN 412.
FT REPEAT 14068 14102 PLECTIN 413.
FT REPEAT 14102 14136 PLECTIN 414.
FT REPEAT 14136 14170 PLECTIN 415.
FT REPEAT 14170 14204 PLECTIN 416.
FT REPEAT 14204 14238 PLECTIN 417.
FT REPEAT 14238 14272 PLECTIN 418.
FT REPEAT 14272 14306 PLECTIN 419.
FT REPEAT 14306 14340 PLECTIN 420.
FT REPEAT 14340 14374 PLECTIN 421.
FT REPEAT 14374 14408 PLECTIN 422.
FT REPEAT 14408 14442 PLECTIN 423.
FT REPEAT 14442 14476 PLECTIN 424.
FT REPEAT 14476 14510 PLECTIN 425.
FT REPEAT 14510 14544 PLECTIN 426.
FT REPEAT 14544 14578 PLECTIN 427.
FT REPEAT 14578 14612 PLECTIN 428.
FT REPEAT 14612 14646 PLECTIN 429.
FT REPEAT 14646 14680 PLECTIN 430.
FT REPEAT 14680 14714 PLECTIN 431.
FT REPEAT 14714 14748 PLECTIN 432.
FT REPEAT 14748 14782 PLECTIN 433.
FT REPEAT 14782 14816 PLECTIN 434.
FT REPEAT 14816 14850 PLECTIN 435.
FT REPEAT 14850 14884 PLECTIN 436.
FT REPEAT 14884 14918 PLECTIN 437.
FT REPEAT 14918 14952 PLECTIN 438.
FT REPEAT 14952 14986 PLECTIN 439.
FT REPEAT 14986 15020 PLECTIN 440.
FT REPEAT 15020 15054 PLECTIN 441.
FT REPEAT 15054 15088 PLECTIN 442.
FT REPEAT 15088 15122 PLECTIN 443.
FT REPEAT 15122 15156 PLECTIN 444.
FT REPEAT 15156 15190 PLECTIN 445.
FT REPEAT 15190 15224 PLECTIN 446.
FT REPEAT 15224 15258 PLECTIN 447.
FT REPEAT 15258 15292 PLECTIN 448.
FT REPEAT 15292 15326 PLECTIN 449.
FT REPEAT 15326 15360 PLECTIN 450.
FT REPEAT 15360 15394 PLECTIN 451.
FT REPEAT 15394 15428 PLECTIN 452.
FT REPEAT 15428 15462 PLECTIN 453.
FT REPEAT 15462 15496 PLECTIN 454.
FT REPEAT 15496 15530 PLECTIN 455.
FT REPEAT 15530 15564 PLECTIN 456.
FT REPEAT 15564 15598 PLECTIN 457.
FT REPEAT 15598 15632 PLECTIN 458.
FT REPEAT 15632 15666 PLECTIN 459.
FT REPEAT 15666 15700 PLECTIN 460.
FT REPEAT 15700 15734 PLECTIN 461.
FT REPEAT 15734 15768 PLECTIN 462.
FT REPEAT 15768 15802 PLECTIN 463.
FT REPEAT 15802 15836 PLECTIN 464.
FT REPEAT 15836 15870 PLECTIN 465.
FT REPEAT 15870 15904 PLECTIN 466.
FT REPEAT 15904 15938 PLECTIN 467.
FT REPEAT 15938 15972 PLECTIN 468.
FT REPEAT 15972 16006 PLECTIN 469.
FT REPEAT 16006 16040 PLECTIN 470.
FT REPEAT 16040 16074 PLECTIN 471.
FT REPEAT 16074 16108 PLECTIN 472.
FT REPEAT 16108 16142 PLECTIN 473.
FT REPEAT 16142 16176 PLECTIN 474.
FT REPEAT 16176 16210 PLECTIN 475.
FT REPEAT 16210 16244 PLECTIN 476.
FT REPEAT 16244 16278 PLECTIN 477.
FT REPEAT 16278 16312 PLECTIN 478.
FT REPEAT 16312 16346 PLECTIN 479.
FT REPEAT 16346 16380 PLECTIN 480.
FT REPEAT 16380 16414 PLECTIN 481.
FT REPEAT 16414 16448 PLECTIN 482.
FT REPEAT 16448 16482 PLECTIN 483.
FT REPEAT 16482 16516 PLECTIN 484.
FT REPEAT 16516 16550 PLECTIN 485.
FT REPEAT 16550 16584 PLECTIN 486.
FT REPEAT 16584 16618 PLECTIN 487.
FT REPEAT 16618 16652 PLECTIN 488.
FT REPEAT 16652 16686 PLECTIN 489.
FT REPEAT 16686 16720 PLECTIN 490.
FT REPEAT 16720 16754 PLECTIN 491.
FT REPEAT 16754 16788 PLECTIN 492.
FT REPEAT 16788 16822 PLECTIN 493.
FT REPEAT 16822 16856 PLECTIN 494.
FT REPEAT 16856 16890 PLECTIN 495.
FT REPEAT 16890 16924 PLECTIN 496.
FT REPEAT 16924 16958 PLECTIN 497.
FT REPEAT 16958 16992 PLECTIN 498.
FT REPEAT 16992 17026 PLECTIN 499.
FT REPEAT 17026 17060 PLECTIN 500.
FT REPEAT 17060 17094 PLECTIN 501.
FT REPEAT 17094 17128 PLECTIN 502.
FT REPEAT 17128 17162 PLECTIN 503.
FT REPEAT 17162 17196 PLECTIN 504.
FT REPEAT 17196 17230 PLECTIN 505.
FT REPEAT 17230 17264 PLECTIN 506.
FT REPEAT 17264 17298 PLECTIN 507.
FT REPEAT 17298 17332 PLECTIN 508.
FT REPEAT 17332 17366 PLECTIN 509.
FT REPEAT 17366 17400 PLECTIN 510.
FT REPEAT 17400 17434 PLECTIN 511.
FT REPEAT 17434 17468 PLECTIN 512.
FT REPEAT 17468 17502 PLECTIN 513.
FT REPEAT 17502 17536 PLECTIN 514.
FT REPEAT 17536 17570 PLECTIN 515.
FT REPEAT 17570 17604 PLECTIN 516.
FT REPEAT 17604 17638 PLECTIN 517.
FT REPEAT 17638 17672 PLECTIN 518.
FT REPEAT 17672 17706 PLECTIN 519.
FT REPEAT 17706 17740 PLECTIN 520.
FT REPEAT 17740 17774 PLECTIN 521.
FT REPEAT 17774 17808 PLECTIN 522.
FT REPEAT 17808 17842 PLECTIN 523.
FT REPEAT 17842 17876 PLECTIN 524.
FT REPEAT 17876 17910 PLECTIN 525.
FT REPEAT 17910 17944 PLECTIN 526.
FT REPEAT 17944 17978 PLECTIN 527.
FT REPEAT 17978 18012 PLECTIN 528.
FT REPEAT 18012 18046 PLECTIN 529.
FT REPEAT 18046 18080 PLECTIN 530.
FT REPEAT 18080 18114 PLECTIN 531.
FT REPEAT 18114 18148 PLECTIN 532.
FT REPEAT 18148 18182 PLECTIN 533.
FT REPEAT 18182 18216 PLECTIN 534.
FT REPEAT 18216 
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 14:01:53 ; Search time 17 Seconds
(without alignments)
192.383 Million cell updates/sec

Title: US-10-010-667a-2_COPY_185_218

Perfect score: 1 RRSYRYKLLNWAYQQVQNKEDAWIEHDVWRMEI 34
Sequence:

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	62.5	32.6	480	2 B70446	hypothetical prote
2	58.5	30.5	1418	2 S40764	hypothetical prote
3	58	30.2	900	2 E69631	galactosamine-cont
4	56	29.2	150	2 T28409	ORF MSV248 probabl
5	55.5	28.9	502	2 S49113	hypothetical prote
6	53	27.6	152	2 C69546	hypothetical prote
7	52	27.1	79	2 E69966	phorbol ester-bind
8	52	27.1	1734	2 A41101	hypothetical prote
9	52	27.1	1813	2 T19295	hypothetical prote
10	51.5	26.8	713	2 G4606	hydantoin utilizat
11	51	26.6	90	2 AB2102	hypothetical prote
12	51	26.6	444	2 T27222	hypothetical prote
13	51	26.6	495	1 S75895	probable L-xylosyl
14	51	26.6	793	2 C83260	hypothetical prote
15	50.5	26.3	201	2 AG2434	hypothetical prote
16	50.5	26.3	402	2 AG1839	hypothetical prote
17	50.5	26.3	712	2 G71907	probable hydantoin
18	50	26.0	204	2 AC2411	hypothetical prote
19	50	26.0	784	2 T43510	probable spindle p
20	50	26.0	816	2 T00919	hypothetical prote
21	50	26.0	1575	2 S68448	synaptojanin, 170K
22	50	26.0	1697	2 T00079	hypothetical prote
23	50	26.0	1908	2 T42707	hypothetical prote
24	49.5	25.8	133	2 E84502	hypothetical prote
25	49.5	25.8	224	2 A10697	probable pathogeni
26	49.5	25.8	274	2 H84956	2,3,4,5-tetrahydro
27	49.5	25.8	350	2 H64458	hypothetical prote
28	49.5	25.8	414	2 T16602	hypothetical prote
29	49.5	25.8	787	2 E71537	probable 60kda inn

30 49.5 25.8 940 2 F86502
31 49.5 25.8 940 2 B72120
32 49.5 25.8 1391 2 T20642
33 49.5 25.8 1397 2 B87998
34 49 25.5 257 2 A82691
35 49 25.5 263 2 A82069
36 49 25.5 383 2 H84700
37 49 25.5 388 2 T40609
38 49 25.5 500 2 AE2032
39 49 25.5 1120 2 H88449
40 49 25.5 2925 2 T00133
41 48.5 25.3 266 2 C90257
42 48.5 25.3 304 2 T04751
43 48.5 25.3 339 2 T33477
44 48.5 25.3 505 2 B81021
45 48.5 25.3 638 2 T20944

alyl tRNA syntheta
valine-tRNA ligase
hypothetical prote
protein F09C3.1 [i
tRNA pseudouridine
hypothetical prote
hypothetical prote
hypothetical prote
glycerol kinase [i
protein F54D8.1 [i
RNA-directed RNA p
translation initia
hypothetical prote
hypothetical prote
conserved hypochet
hypothetical prote

ALIGNMENTS

RESULT 1

B70446

hypothetical protein aq_1687 - Aquifex aeolicus

C/Species: Aquifex aeolicus

C/Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999

C/Accession: B70446

R/Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O. V.

Nature 392, 353-358, 1998

A/Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A/Reference number: A70300; MUID:98196666; PMID:9537320

A/Accession: B70446

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-480 <AQF>

A/Cross-references: GB:AE000751; NID:g2984013; PIDN:AAC07560.1; PID:g2984016; GB:AE00066

A/Experimental source: strain VF5

C/Genetics:

A/Gene: aq_1687

Query Match 32.6% Score 62.5; DB 2; Length 480;

Best Local Similarity 25.0%; Pred. No. 1.8; Mismatches 4; Indels 15; Gaps 2;

Matches 12; Conservative 17; Mismatches 4; Indels 15; Gaps 2;

Qy 1 RRSYRYKLLNWAYQ--QVQNKEDAWIEHDVWRME 33

Db 10 RKTLYLKIPAKVSDGKEYKVLDSYEGFRIEKSKEDVFEDKVKYKVK 57

RESULT 2

S40764

hypothetical protein ZK512.3 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 24-Nov-1999

C/Accession: S40764

R/Hawkins, T.; Ainscough, R.

A/Reference number: S40759

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-1418 <HAW>

A/Cross-references: EMBL:Z22177; NID:g297989; PID:g297992

C/Genetics:

C/Introns: 199/1; 238/1; 290/2; 529/3; 557/3; 588/3; 677/2; 733/3; 772/3; 846/3; 946/1;

C/Superfamily: Caenorhabditis elegans hypothetical protein ZK512.3

Query Match 30.5%; Score 58.5; DB 2; Length 1418;

Best Local Similarity 43.5%; Pred. No. 19;

Matches 10; Conservative 5; Mismatches 5; Indels 3; Gaps 1;

Qy 15 QVQNKED---AWIEHDVWRMEI 34

```

hypothetical protein 2 - Microcystis aeruginosa
C/Species: Microcystis aeruginosa
C/Date: 01-Feb-1995 #sequence_revision 12-May-1995 #text_change 23-Mar-2001
C/Accession: S49113
R/Juerchott, K.; Boerner, T.
submitted to the EMBL Data Library, November 1993
A/Description: Sequence of the cyanobacterial plasmid pMA1 from Microcystis aeruginosa
A/Reference number: S49112
A/Accession: S49113
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-502 <JUE>
A/Cross-references: EMBL:Z28337; NID:g509352; PIDN:CAA82191.1; PID:g509354
C/Superfamily: Microcystis aeruginosa hypothetical protein 2

Query Match      28.9%; Score 55.5; DB 2; Length 502;
Best Local Similarity 25.0%; Pred. No. 16;
Matches 12; Conservative 7; Mismatches 14; Indels 15; Gaps 1;

QY      2 RSYRYKLNNLA-----YQQVQONKEDAWIEHDVWMEI 34
      :|||:|||
Db      392 QAVREKLLGHAVWRYLDGECQLPEEFKALQAEANKQWENNDSEDEL 439

RESULT 6
C69546
hypothetical protein AF2371 - Archaeoglobus fulgidus
C/Species: Archaeoglobus fulgidus
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C/Accession: C69546
R/Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

```

A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Attiach, P.; Kaine, B.P.; Sykes, S.M.; Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeal
A:Reference number: A69250; MUID:98049343; PMID:9389475
A:Accession: C69546
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 152,478

```

Query Match          27.6%; Score 53; DB 2; Length 152;
Best Local Similarity 42.9%; Pred. No. 9.4;
Matches 9; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY      5 RYKLLNWAYQQVQCNKEDAWI 25
          |||||:|:|:|:|:|
Db      21 RYKLLSGLVLDYERNRVGAYI 41

RESULT 7
E69966
hypothetical protein yqkC - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
C:Accession: E69966
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

```

A.A.: Ehrlich, S.D.; Emmerson, P.T.; Etian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997

AA:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizi, A.; Galle
iech, J.; Harwood, C.R.; Henauf, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koepter, P.; Konigstein, G.; Krogh, S.; Kumano, K.; Kurita, K.; Lapidus, A.; Lardiniois
A., Authors: Lauber, J.; Lazarevic, V.; Lee, S.; Levine, A.; Liu, H.; Masuda, S.; Mauee
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl T.M.; Portetel
M.

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Togmoni, A.; Tosato, V.; Uchiyama, T.; Winterts, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, R.; Ruthven, P.; Vachon, W.

A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: E69966
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-79 <KUN>
A;Cross-references: GB:Z99116; GB:AL009126; NID:G2634723; PIDN:CAB14297.1; PID:e1185634;
A;Experimental source: strain 168
C;Genetics:
A;Gene: yqkC

Query Match 27.1%; Score 52; DB 2; Length 79;
Best Local Similarity 28.4%; Pred. No. 6.4;
Matches 9; Conservative 8; Mismatches 15; Indels 0; Gaps 0;

QY 2 RSYRYKLHWAYQQVQCNKEDAWIEHDVWRME 33
::: |||::||::||::||::||::||::||::||
Db 23 QNHPTLLHWSIGAESIKKDVLQLQDETTFE 54

RESULT 8
A41101
phorbol ester-binding protein unc-13 - *Caenorhabditis elegans*
C;Species: *Caenorhabditis elegans*
C;Date: 20-Mar-1992 #sequence_revision 20-Mar-1992 #text_change 23-Feb-1997
C;Accession: A41101
R;Maruyama, I.N.; Brenner, S.
Proc. Natl. Acad. Sci. U.S.A. 88, 5729-5733, 1991
A;Title: A phorbol ester/diacylglycerol-binding protein encoded by the unc-13 gene of *Caenorhabditis elegans*
A;Reference number: A41101; MUID:91288538; PMID:2062851
A;Accession: A41101
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1734 <MAR>
A;Cross-references: GB:M62830
C;Superfamily: protein kinase C zinc-binding repeat homology
C;Keywords: phosphoprotein
F;615-664/Domain: protein kinase C zinc-binding repeat homology <KZ2>

Query Match 27.1%; Score 52; DB 2; Length 1734;
Best Local Similarity 40.0%; Pred. No. 1.7e+02;
Matches 8; Conservative 5; Mismatches 5; Indels 2; Gaps 1;

QY 10 NWAYQQVQQ--NKEDAWIEH 27
|||::||::||::||::||::||::||::||
Db 430 NWRDYSIQEDNEKDNWKQH 449

RESULT 9
T19295
hypothetical protein ZK524.2a - *Caenorhabditis elegans*
C;Species: *Caenorhabditis elegans*
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 02-Sep-2000
C;Accession: T19295; T19931; T27889
R;Gardner, A.
submitted to the EMBL Data Library, September 1996
A;Reference number: Z19103
A;Accession: T19295
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1813 <WIL>
A;Cross-references: EMBL:Z79694; PIDN:CAB01966.1; GSFP:GN00019; CESP:ZK524.2a
A;Experimental source: Clone C15A11
R;Lloyd, C.
submitted to the EMBL Data Library, March 1997
A;Reference number: Z19198
A;Accession: T19931
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1813 <WT2>
A;Cross-references: EMBL:Z92779; PIDN:CAB07173.1; GSFP:GN00019; CESP:ZK524.2a
A;Experimental source: clone C44E1
R;Gardner, A.

RESULT 14

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 14:10:33 ; Search time 60.093 Seconds
(without alignments)
176.480 Million cell updates/sec

Title: US-10-010-667A-2_COPY_185_218

Perfect score: 192

Sequence: 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 34

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1279676 seqs, 311918243 residues

Total number of hits satisfying chosen parameters: 1279676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

```
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	192	100.0	34	14	US-10-011-095-20
2	192	100.0	34	14	US-10-010-667A-20
3	192	100.0	34	14	US-10-165-044-38
4	192	100.0	267	9	US-09-747-835A-50
5	192	100.0	267	12	US-09-747-835A-51
6	192	100.0	267	12	US-10-312-312-50
7	192	100.0	267	12	US-10-312-312-51
8	192	100.0	339	9	US-09-759-143-879
9	192	100.0	339	9	US-09-780-669-879
10	192	100.0	339	9	US-09-822-827-879
11	192	100.0	339	9	US-09-802-520-11
12	192	100.0	339	9	US-09-895-793-879
13	192	100.0	339	9	US-09-895-814-879
14	192	100.0	339	12	US-10-205-267-13
15	192	100.0	339	12	US-10-408-009-2

16	192	100.0	339	13	US-10-012-896-879
17	192	100.0	339	14	US-10-011-095-2
18	192	100.0	339	14	US-10-010-667A-2
19	192	100.0	339	14	US-10-205-823-397
20	192	100.0	339	14	US-10-144-678A-879
21	192	100.0	339	14	US-10-294-025-879
22	192	100.0	339	15	US-10-239-607-37
23	192	100.0	339	15	US-10-295-027-714
24	192	100.0	339	15	US-10-295-027-1347
25	192	100.0	368	12	US-10-425-114-72779
26	192	100.0	375	14	US-10-165-044-2
27	110	57.3	443	12	US-10-455-822-93
28	107	55.7	141	9	US-09-963-896-1
29	107	55.7	173	14	US-10-011-095-8
30	107	55.7	173	14	US-10-010-667A-8
31	107	55.7	237	9	US-09-747-835A-15
32	107	55.7	237	12	US-10-312-312-15
33	107	55.7	419	12	US-10-455-822-11
34	107	55.7	419	12	US-10-455-822-80
35	107	55.7	419	12	US-10-455-822-172
36	107	55.7	419	12	US-10-455-822-174
37	107	55.7	419	12	US-10-455-822-195
38	107	55.7	419	15	US-10-239-607-32
39	107	55.7	444	12	US-10-455-822-88
40	107	55.7	444	12	US-10-455-822-89
41	107	55.7	444	12	US-10-455-822-90
42	107	55.7	444	12	US-10-455-822-91
43	107	55.7	444	12	US-10-455-822-92
44	107	55.7	454	10	US-09-888-257A-10
45	107	55.7	454	12	US-10-455-822-3

ALIGNMENTS

RESULT 1

```
US-10-011-095-20
; Sequence 20, Application US/10011095
; Publication No. US20030045682A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Hubert, Rene S.
; APPLICANT: Leong, Kahan
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Safran, Douglas C.
; APPLICANT: Mitchell, Steve Chappell
; TITLE OF INVENTION: ANTIBODIES IMMUNOSPECIFIC FOR STEAP1 (AS AMENDED)
; FILE REFERENCE: 511582001610
; CURRENT APPLICATION NUMBER: US/10/011.095
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: 09/323,873
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: STEAP-1 peptide
US-10-011-095-20
```

Query Match 100.0%; Score 192; DB 14; Length 34;

Best Local Similarity 100.0%; Pred. No. 1.8e-18;

Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 34

Db 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 34

```
RESULT 2
US-10-010-667A-20
; Sequence 20, Application US/10010667A
; Publication No. US20030055217A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Hubert, Rene S.
; APPLICANT: Leong, Kahan
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Saffran, Douglas C.
; APPLICANT: Mitchell, Steve Chappell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USES THEREOF
; FILE REFERENCE: 511582001601
; CURRENT APPLICATION NUMBER: US/10/010,667A
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: 09/323,873
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: STEAP-1 peptide
US-10-010-667A-20

Query Match      100.0%; Score 192; DB 14; Length 34;
Best Local Similarity 100.0%; Pred. No. 1.8e-18;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSYRYKLLNWAYQQVQQNKEDAWIEHDVWRMEI 34
Db 1 RRSYRYKLLNWAYQQVQQNKEDAWIEHDVWRMEI 34

RESULT 3
US-10-165-044-38
; Sequence 38, Application US/10165044
; Publication No. US20030149531A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Rene S. Hubert
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas Saffran
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Steven Chappell Mitchell
; APPLICANT: Mary Faris
; APPLICANT: Ava Jakobovits
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USES THEREOF
; FILE REFERENCE: 51158-20016.02
; CURRENT APPLICATION NUMBER: US/10/165,044
; CURRENT FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: US 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: US 60/091,183
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 09/323,873
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: US 09/455,486
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: WO 99/62941
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/33040
; PRIOR FILING DATE: 2000-12-06
```

```
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-165-044-38

Query Match      100.0%; Score 192; DB 14; Length 34;
Best Local Similarity 100.0%; Pred. No. 1.8e-18;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSYRYKLLNWAYQQVQQNKEDAWIEHDVWRMEI 34
Db 1 RRSYRYKLLNWAYQQVQQNKEDAWIEHDVWRMEI 34

RESULT 4
US-09-747-835A-50
; Sequence 50, Application US/09747835A
; Patent No. US20020146692A1
; GENERAL INFORMATION:
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Asundi, Vinod
; APPLICANT: Drmauac, Radoje T
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIKE
; TITLE OF INVENTION: LIKE) POLYPEPTIDES AND POLYNUCLEOTIDES
; FILE REFERENCE: HYS-37CIP
; CURRENT APPLICATION NUMBER: US/09/747,835A
; CURRENT FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: US 09/729,739
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: US 09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 50
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-747-835A-50

Query Match      100.0%; Score 192; DB 9; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.6e-17;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSYRYKLLNWAYQQVQQNKEDAWIEHDVWRMEI 34
Db 119 RRSYRYKLLNWAYQQVQQNKEDAWIEHDVWRMEI 152

RESULT 5
US-09-747-835A-51
; Sequence 51, Application US/09747835A
; Patent No. US20020146692A1
; GENERAL INFORMATION:
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
```

```
; APPLICANT: Zhou, Ping
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Asundi, Vinod
; APPLICANT: Drmanac, Radoje T
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIKE
; TITLE OF INVENTION: LIKE) POLYPEPTIDES AND POLYNUCLEOTIDES
; FILE REFERENCE: HYS-37CIP
; CURRENT APPLICATION NUMBER: US/09/747,835A
; CURRENT FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: US 09/729,739
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: US 09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 51
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-747-835A-51
```

```
Query Match 100.0%; Score 192; DB 9; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.6e-17;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVVRMEI 34
Db 119 RRSRYKLLNWAYQQVQONKEDAWIEHDVVRMEI 152
```

RESULT 6

```
US-10-312-312-50
; Sequence 50, Application US/10312312
; Publication No. US20040068097A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Asundi, Vinod
; APPLICANT: Drmanac, Radoje T
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIKE
; TITLE OF INVENTION: LIKE) POLYPEPTIDES AND POLYNUCLEOTIDES
; FILE REFERENCE: 21272-015-061/HYS-37CIP
; CURRENT APPLICATION NUMBER: US/10/312,312
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 09/729,739
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: US 09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 63
```

```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 50
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-312-312-50
```

```
Query Match 100.0%; Score 192; DB 12; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.6e-17;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVVRMEI 34
Db 119 RRSRYKLLNWAYQQVQONKEDAWIEHDVVRMEI 152
```

RESULT 7

```
US-10-312-312-51
; Sequence 51, Application US/10312312
; Publication No. US20040068097A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Asundi, Vinod
; APPLICANT: Drmanac, Radoje T
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIKE
; TITLE OF INVENTION: LIKE) POLYPEPTIDES AND POLYNUCLEOTIDES
; FILE REFERENCE: 21272-015-061/HYS-37CIP
; CURRENT APPLICATION NUMBER: US/10/312,312
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 09/729,739
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: US 09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 51
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-312-312-51
```

```
Query Match 100.0%; Score 192; DB 12; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.6e-17;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVVRMEI 34
Db 119 RRSRYKLLNWAYQQVQONKEDAWIEHDVVRMEI 152
```

RESULT 8

```
US-09-759-143-879
; Sequence 879, Application US/09759143
; Patent No. US2002002248A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
```

; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C23
; CURRENT APPLICATION NUMBER: US/09/759,143
; CURRENT FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-759-143-879

Query Match 100.0%; Score 192; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-17;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 34
|||||
DB 185 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 218

RESULT 9

US-09-780-669-879
; Sequence 879, Application US/09780669
; Patent No. US20020051977A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C24
; CURRENT APPLICATION NUMBER: US/09/780,669
; CURRENT FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 943
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-780-669-879

Query Match 100.0%; Score 192; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-17;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 34
|||||
DB 185 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 218

RESULT 12

US-09-895-793-879

Query Match 100.0%; Score 192; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-17;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 34
|||||
DB 185 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 218

RESULT 10

US-09-822-827-879
; Sequence 879, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C1
; CURRENT APPLICATION NUMBER: US/09/822,827
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-822-827-879

Query Match 100.0%; Score 192; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-17;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 34
|||||
DB 185 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 218

RESULT 11

US-09-802-520-11
; Sequence 11, Application US/09802520
; Publication No. US20020187472A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Paris, Mary
; APPLICANT: Chen, Huei-Mei
; APPLICANT: Ison, Craig H.
; TITLE OF INVENTION: STRAP-RELATED PROTEIN
; FILE REFERENCE: PC-0037 US
; CURRENT APPLICATION NUMBER: US/09/802,520
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PERL Program
; SEQ ID NO 11
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020187472A1 g6572948
US-09-802-520-11

Query Match 100.0%; Score 192; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-17;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 34
|||||
DB 185 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 218

; Sequence 879, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121-534C2
; CURRENT APPLICATION NUMBER: US/09/895,793
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-895-793-879

Query Match 100.0%; Score 192; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-17;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVVRMEI 34
|||||
Db 185 RRSRYKLLNWAYQQVQONKEDAWIEHDVVRMEI 218

RESULT 13
US-09-895-814-879
; Sequence 879, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa

; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121-427C26
; CURRENT APPLICATION NUMBER: US/09/895,814
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-895-814-879

Query Match 100.0%; Score 192; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-17;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVVRMEI 34
|||||
Db 185 RRSRYKLLNWAYQQVQONKEDAWIEHDVVRMEI 218

RESULT 14
US-10-205-267-13
; Sequence 13, Application US/10205267
; Publication No. US20030064397A1
; GENERAL INFORMATION:
; APPLICANT: Spaccake, Kimberly M.
; APPLICANT: Rickert Paula K.
; APPLICANT: Lal, Preeti G.
; APPLICANT: Ison, Craig H.
; TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN PROSTATE AND LI
; TITLE OF INVENTION: TUMORS
; FILE REFERENCE: PV-0008 CIP
; CURRENT APPLICATION NUMBER: US/10/205,267
; CURRENT FILING DATE: 2002-07-24
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PERL Program
; SEQ ID NO 13
; LENGTH: 339
; TYPE: PRT
; ORGANISM:
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: Genbank ID No: 96572948
US-10-205-267-13

Query Match 100.0%; Score 192; DB 12; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-17;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVVRMEI 34
|||||
Db 185 RRSRYKLLNWAYQQVQONKEDAWIEHDVVRMEI 218

RESULT 15
US-10-408-009-2
; Sequence 2, Application US/10408009
; Publication No. US20040072196A1
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Stephen C. Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USES THEREOF
; FILE REFERENCE: 511582001603
; CURRENT APPLICATION NUMBER: US/10/408,009
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 09/455,486
; PRIOR FILING DATE: 1999-12-06

```

; PRIOR APPLICATION NUMBER: 09/323,873
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-009-2

Query Match      100.0%; Score 192; DB 12; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-17;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RRSYRYKLLNWAYQQVQQNKEDAWIEHDVWRMEI 34
        |||||||
Db      185 RRSYRYKLLNWAYQQVQQNKEDAWIEHDVWRMEI 218

Search completed: July 12, 2004, 14:27:49
Job time : 60.093 secs

```


GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 14:04:08 ; Search time 20.9535 Seconds
(without alignments)
83.770 Million cell updates/sec

Title: US-10-010-667A-2_COPY_185_218

Perfect score: 192

Sequence: 1 RRSRYKLLNWAYQQVQNKEDAWIEHDVWRMEI 34

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTUS COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/PTUS COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	192	100.0	34	4	US-09-323-873A-20
2	192	100.0	339	4	US-09-323-873A-2
3	192	100.0	339	4	US-09-685-166A-879
4	107	55.7	141	3	US-09-083-521-1
5	107	55.7	173	4	US-09-323-873A-8
6	85	44.3	15	4	US-09-323-873A-32
7	57	29.7	320	4	US-09-543-681A-5163
8	52.5	27.3	340	4	US-09-543-681A-7850
9	52	27.1	166	4	US-09-328-352-4565
10	51	26.6	789	4	US-09-252-991A-28767
11	50	26.0	370	4	US-09-149-476-436
12	49.5	25.8	940	4	US-09-198-452A-111
13	49	25.5	30	4	US-09-205-258-646
14	49	25.5	120	4	US-09-205-258-644
15	48.5	25.3	654	3	US-09-560-005-10
16	48.5	25.3	654	3	US-09-418-540-10
17	48.5	25.3	654	4	US-09-969-528-10
18	48	25.0	405	4	US-09-134-000C-5472
19	48	25.0	2037	4	US-09-543-681A-5538
20	47.5	24.7	178	4	US-09-134-001C-3412
21	47.5	24.7	4150	3	US-09-428-517-2
22	47	24.5	258	4	US-09-107-532A-6289
23	47	24.5	644	4	US-09-198-452A-63
24	47	24.5	1503	4	US-09-600-087-2
25	46.5	24.2	379	4	US-09-252-991A-26357
26	46.5	24.2	602	4	US-09-489-039A-12436
27	46.5	24.2	927	4	US-09-540-236-2607

28	46	24.0	111	4	US-09-107-532A-5449	Sequence 5449, Ap
29	46	24.0	236	3	US-09-121-973-4	Sequence 4, Appli
30	46	24.0	236	3	US-09-332-319-4	Sequence 4, Appli
31	46	24.0	236	4	US-09-239-867-2	Sequence 2, Appli
32	46	24.0	477	4	US-09-198-452A-19	Sequence 19, Appl
33	46	24.0	599	4	US-09-489-039A-13050	Sequence 13050, A
34	46	24.0	605	4	US-09-252-991A-26916	Sequence 26916, A
35	46	24.0	961	4	US-09-328-352-7358	Sequence 7358, Ap
36	45.5	23.7	441	4	US-09-107-532A-4173	Sequence 4173, Ap
37	45.5	23.7	605	4	US-09-252-991A-20980	Sequence 20980, A
38	45.5	23.7	854	4	US-09-206-551-17	Sequence 17, Appl
39	45	23.4	280	4	US-09-489-039A-9395	Sequence 9395, Ap
40	45	23.4	371	1	US-08-746-789A-2	Sequence 2, Appli
41	45	23.4	371	4	US-09-570-593-5	Sequence 5, Appli
42	45	23.4	516	4	US-09-489-039A-9477	Sequence 9477, Ap
43	45	23.4	526	4	US-09-198-452A-67	Sequence 67, Appl
44	45	23.4	631	4	US-09-107-532A-6640	Sequence 6640, Ap
45	45	23.4	837	4	US-09-489-039A-9385	Sequence 9385, Ap

ALIGNMENTS

RESULT 1
US-09-323-873A-20
; Sequence 20, Application US/09323873A
; Patent No. 6329503
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Kahan Leong
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Steve Chappell Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; FILE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USES THEREOF
; FILE REFERENCE: 129.16USU2
; CURRENT APPLICATION NUMBER: US/09/323,873A
; CURRENT FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: STEAP-1 PEPTIDE
US-09-323-873A-20

Query Match 100.0%; Score 192; DB 4; Length 34;
Best Local Similarity 100.0%; Pred. No. 5.5e-20;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RRSRYKLLNWAYQQVQNKEDAWIEHDVWRMEI 34
Db 1 RRSRYKLLNWAYQQVQNKEDAWIEHDVWRMEI 34

RESULT 2
US-09-323-873A-2
; Sequence 2, Application US/09323873A
; Patent No. 6329503
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Kahan Leong
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Steve Chappell Mitchell

; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; FILE REFERENCE: EXPRESSED IN HUMAN CANCERS AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/323,873A
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-323-873A-2

Query Match 100.0%; Score 192; DB 4; Length 339;
Best Local Similarity 100.0%; Pred. No. 7.2e-19;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSRYKLLNWAYQQVQNKEDAWIEHDVWRMEI 34
(((|||||))))
Db 185 RRSRYKLLNWAYQQVQNKEDAWIEHDVWRMEI 218

RESULT 3
US-09-685-166A-879
; Sequence 879, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedrick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685,166A
; CURRENT FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-685-166A-879

Query Match 100.0%; Score 192; DB 4; Length 339;
Best Local Similarity 100.0%; Pred. No. 7.2e-19;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSRYKLLNWAYQQVQNKEDAWIEHDVWRMEI 34
(((|||||))))
Db 185 RRSRYKLLNWAYQQVQNKEDAWIEHDVWRMEI 218

RESULT 4
US-09-083-521-1
; Sequence 1, Application US/09083521

; Patent No. 6048970
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: PROSTATE GROWTH-ASSOCIATED MEMBRANE PROTEINS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/083,521
; FILING DATE: Herewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0527 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 141 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PROSTUT10
; CLONE: 1691243
US-09-083-521-1

Query Match 55.7%; Score 107; DB 3; Length 141;
Best Local Similarity 58.8%; Pred. No. 1.5e-07;
Matches 20; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 RRSRYKLLNWAYQQVQNKEDAWIEHDVWRMEI 34
|||||
Db 13 RRSRYKLLNWAYQQVQHANIENSWNEEVWRIEM 46

RESULT 5
US-09-323-873A-8
; Sequence 8, Application US/09323873A
; Patent No. 6329503
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Kahan Leong
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Steve Chappell Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; FILE REFERENCE: 129.16USU2
; CURRENT APPLICATION NUMBER: US/09/323,873A
; CURRENT FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8

; LENGTH: 173
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-323-873A-8

Query Match 55.7%; Score 107; DB 4; Length 173;
Best Local Similarity 58.8%; Pred. No. 1.9e-07;
Matches 20; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 RBSRYKLNWAYQQVQQNKEDAWIEHDVWRMEI 34
DB 81 RRSERYFLNWAYQQVQQHANIENSWNEEVWRMEI 114

RESULT 6

US-09-323-873A-32
; Sequence 32, Application US/09323873A

; Patent No. 6329503
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Kahan Leong
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Steve Chappell Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USES THEREOF
; FILE REFERENCE: 129.16USU2
; CURRENT APPLICATION NUMBER: US/09/323,873A

; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 32
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo Sapiens

US-09-323-873A-32

Query Match 44.3%; Score 85; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 YQVQVQNKEDAWIEH 27
DB 1 YQVQVQNKEDAWIEH 15

RESULT 7

US-09-543-681A-5163
; Sequence 5163, Application US/09543681A

; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344

; SEQ ID NO 5163
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5163

Query Match 29.7%; Score 57; DB 4; Length 320;
Best Local Similarity 38.7%; Pred. No. 3.2;

Matches 12; Conservative 5; Mismatches 8; Indels 6; Gaps 2;

QY 6 YKLLN-WAYQQVQQNKED-----AWIEHDVW 30
DB 78 YAVMNEWAYNEYNQKDGGLSAAWGVNDRW 108

RESULT 8

US-09-543-681A-7850
; Sequence 7850, Application US/09543681A

; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05

; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7850
; LENGTH: 340
; TYPE: PRT

; ORGANISM: Proteus mirabilis
US-09-543-681A-7850

Query Match 27.3%; Score 52.5; DB 4; Length 340;
Best Local Similarity 30.3%; Pred. No. 14;
Matches 10; Conservative 8; Mismatches 8; Indels 7; Gaps 1;

QY 3 SYRY-----KLLNWAYQQVQQNKEDAWIEHD 28
DB 169 THYTAGTTDIALKRAYQQMKTLEEWLKR 201

RESULT 9

US-09-328-352-4565
; Sequence 4565, Application US/09328352

; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA

; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252

; SEQ ID NO 4565
; LENGTH: 166
; TYPE: PRT

; ORGANISM: Acinetobacter baumannii
US-09-328-352-4565

Query Match 27.1%; Score 52; DB 4; Length 166;
Best Local Similarity 27.6%; Pred. No. 7.5;
Matches 8; Conservative 7; Mismatches 10; Indels 4; Gaps 1;

QY 4 YRYK----LLNWAYQQVQQNKEDAWIEHD 28
DB 133 YEYNEKGVLLWTHHDFQNRHENGWLKHN 161

RESULT 10

US-09-252-991A-28767

; Sequence 28767, Application US/09252991A
; Patent No. 6551795

; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A

```
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28767
; LENGTH: 789
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28767

Query Match      26.6%; Score 51; DB 4; Length 789;
Best Local Similarity 52.9%; Pred. No. 59;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY      16 VQONKEDAWIEHDVWRM 32
Db      397 VQSKDDAVREHFWRL 413

RESULT 11
US-09-149-476-436
; Sequence 436, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,879
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,880
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,894
```

```

; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

Query Match      26.0%; Score 50; DB 4; Length 370;
Best Local Similarity 40.7%; Pred. No. 35;
Matches 11; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY      5 RYKLLNWAYQQVQONKEDAWIEHDVWR 31
DB      191 RQKLNLDALRLLEELKEFANFDVWR 217
      | : | | | : | | | : | | |
RESULT 12
US-09-198-452A-111
; Sequence 111, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffois, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 111
; LENGTH: 940
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-111

Query Match      25.8%; Score 49.5; DB 4; Length 940;
Best Local Similarity 31.6%; Pred. No. 1.2e+02;
Matches 12; Conservative 8; Mismatches 9; Indels 9; Gaps 2

QY      4 YR-YKLLNW-----AYQQVQONKEDAWIEHDVWRM 32
DB      172 YRGYLVNWDPLVQLTALADDEVEYEEKDGLVIYIRYM 209
      | | | | | | | | | | : | : | : | : | : |

RESULT 13
US-09-205-258-646
; Sequence 646, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: PZ007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894

```

; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,971
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,964
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,882
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,899
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,893
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,900
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,901
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,892
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,915
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/049,019
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,970
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,972
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,916
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/049,373
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,875
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/049,374
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,917
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,949
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,974
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,883
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,897
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,898
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,962
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,963
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,877
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,878
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/070,923
 ; EARLIER FILING DATE: 1997-12-18
 ; EARLIER APPLICATION NUMBER: 60/092,921
 ; EARLIER FILING DATE: 1998-07-15
 ; EARLIER APPLICATION NUMBER: 60/094,657
 ; EARLIER FILING DATE: 1998-07-30
 ; NUMBER OF SEQ ID NOS: 1227
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 646
 ; LENGTH: 30
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-205-258-646

Query Match 25.58; Score 49; DB 4; Length 30;
 Best Local Similarity 38.1%; Pred. No. 2.9;
 Matches 8; Conservative 5; Mismatches 6; Indels 1;

QY 9 LNWAYQ--VQOQNKEDAWIWH 27
 : ||:: ||||: : |

Db 7 VWMFESLQVQNRPERWASH 27
 RESULT 14
 US-09-205-258-644
 ; Sequence 644, Application US/09205258
 ; Patent No. 6525174
 ; GENERAL INFORMATION:
 ; APPLICANT: Young et al.
 ; TITLE OF INVENTION: 207 Human Secreted Proteins
 ; FILE REFERENCE: PZ007P1
 ; CURRENT APPLICATION NUMBER: US/09/205,258
 ; CURRENT FILING DATE: 1998-12-04
 ; EARLIER APPLICATION NUMBER: PCT/US98/11422
 ; EARLIER FILING DATE: 1998-06-04
 ; EARLIER APPLICATION NUMBER: 60/048,885
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/049,375
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,881
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,880
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,896
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/049,020
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,876
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,895
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,884
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,894
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,971
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,964
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,882
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,899
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,893
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,900
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,901
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,892
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,915
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/049,019
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,970
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,972
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,916
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/049,373
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,875
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/049,374
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,917
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,949
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,974
 ; EARLIER FILING DATE: 1997-06-06

```
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 644
; LENGTH: 120
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (105)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (115)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;
US-09-205-258-644

Query Match          25.5%; Score 49; DB 4; Length 120;
Best Local Similarity 38.1%; Pred. No. 14;
Matches 8; Conservative 5; Mismatches 6; Indels 2; Gaps 1;

Qy      9 LNWAYQ--VQONKEDAWIEH 27
       :|::|::|::|::|::|::|
Db      79 VRWAFESLQVPQRPWRWASH 99

RESULT 15
US-08-560-005-10
; Sequence 10, Application US/08560005
; Patent No. 6001354
; GENERAL INFORMATION:
; APPLICANT: Pot, David A.
; APPLICANT: Williams, Lewis T.
; APPLICANT: Jefferson, Anne Bennett
; APPLICANT: Majerus, Philip W.
; TITLE OF INVENTION: No. 6001354el Grb2 Associating Protein and Nucleic
; TITLE OF INVENTION: Acids Encoding Thereof
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,005
; FILING DATE:
; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 2307K-0624000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 654 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Region
; LOCATION: 1..654
; OTHER INFORMATION: /note= "consensus"
;
US-08-560-005-10

Query Match          25.3%; Score 48.5; DB 3; Length 654;
Best Local Similarity 25.0%; Pred. No. 1.1e+02;
Matches 8; Conservative 12; Mismatches 9; Indels 3; Gaps 1;

Qy      3 SYRYKLLNWAYQO---VQONKEDAWIEHDVWR 31
       :|::|::|::|::|::|::|::|::|::|
Db      348 TYKIRGRDYAYKQKDTSEKTRVPAWCDRLWK 379

Search completed: July 12, 2004, 14:12:24
Job time : 21.9535 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 13:51:07 ; Search time 74.7209 Seconds

(without alignments)
128.567 Million cell updates/sec

Title: US-10-010-667A-2_COPY_185_218

Perfect score: 192

Sequence: 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 34

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	192	100.0	34	3	AAY58199 Human STR
2	192	100.0	34	4	Aae02787 Extracell
3	192	100.0	129	4	Aab75315 Human sec
4	192	100.0	129	4	Aab75314 Gene 20 h
5	192	100.0	254	6	Abu98426 STEAP-1 v
6	192	100.0	254	6	Abu98425 STEAP-1 v
7	192	100.0	255	6	Abu98429 STEAP-1 v
8	192	100.0	255	6	Abu98428 STEAP-1 v
9	192	100.0	258	6	Abu98395 Novel hum
10	192	100.0	258	6	Abu98389 Novel hum
11	192	100.0	258	6	Abu98424 STEAP-1 v
12	192	100.0	258	6	Abu98392 Novel hum
13	192	100.0	258	6	Abu98387 Novel hum
14	192	100.0	258	6	Abu98393 Novel hum
15	192	100.0	258	6	Abu98390 Novel hum
16	192	100.0	258	6	Abu98391 Novel hum
17	192	100.0	258	6	Abu98384 Novel hum
18	192	100.0	258	6	Abu98431 STEAP-1 v
19	192	100.0	258	6	Abu98396 Novel hum
20	192	100.0	258	6	Abu98397 Novel hum
21	192	100.0	258	6	Abu98394 Novel hum
22	192	100.0	258	6	Abu98388 Novel hum
23	192	100.0	258	6	Abu98386 Novel hum
24	192	100.0	258	6	Abu98399 Novel hum
25	192	100.0	258	6	Abu98398 Novel hum

ALIGNMENTS

RESULT 1
AAY58199

ID AAY58199 standard; peptide; 34 AA.

XX AAY58199;

XX 14-MAR-2000 (first entry)

XX Human STRAP-1 peptide, corresponding to STRAP-1 extracellular region 2.

XX Serpentine transmembrane antigen of the prostate; STRAP-1; prostate;
transmembrane domain; type IIa membrane protein; expression; cancer;
prostate cancer; bladder cancer; colon cancer; pancreatic cancer;
ovarian cancer; tumour antigen; immunisation; immune response; cellular;
humoral; anticancer vaccine; antibody; detection; diagnosis; prognosis;
monitoring; susceptibility; therapeutic inhibitor; drug targeting;
recombinant protein.

XX Synthetic.

OS Homo sapiens.

PN WO9962941-A2.

XX 09-DEC-1999.

XX 01-JUN-1999; 99WO-US012157.

XX 01-JUN-1998; 98US-0087520P.

XX 30-JUN-1998; 98US-0091183P.

XX (UROC-) UROGENESYS INC.

PA (AFAR/) AFAR D E.

PA (HUBE/) HUBERT R S.

PA (LEON/) LEONG K.

PA (RAIT/) RAITANO A B.

PA (SAFF/) SAFFRAN D C.

XX Afar DE, Hubert RS, Leong K, Raitano AB, Saffran DC;

XX WPI; 2000-072832/06.

XX Novel proteins useful as diagnostic markers and therapeutic targets,
particularly for prostatic cancer.

PS Disclosure; Page 22; 83pp; English.

XX Sequences AAY58198-Y58200 represent synthetic peptides that correspond to
CC the extracellular regions of STRAP-1 (serpentine transmembrane antigen of

Abu60886 Human G p
Abu60887 Human G p
Abu98432 STEAP-1 v
Abu98427 STEAP-1 v
Abu98385 Novel hum
Aay58194 Human STR
Aam01282 p789p ami
Aau69927 Human pro
Aam78845 Human pro
Abu71818 Prostate
Abb95387 Human p78
Abq61813 Prostate
Abu98383 Novel hum
Abu98414 STEAP-1 v
Abu98430 STEAP-1 v
Abr54499 Prostate
Abu63313 Human six
Adb75573 Prostate
Adb14329 Human pro
Aae02780 Human six

the prostate, AAV58194). These peptides were used to raise monoclonal anti-STRAP-1 antibodies. STRAP-1 is the prototype member of the STRAP family of proteins (AAV58194-Y58197) which exhibit a high degree of structural conservation, but which show no significant structural homology to known human proteins. The STRAP-1 gene has been localised to chromosome 7p22. STRAP-1 is thought to be a type IIIa membrane protein and is expressed predominantly in prostatic cells in normal human tissues. Structurally, STRAP-1 is a 339 amino acid protein characterised by six transmembrane domains and intracellular N- and C-termini, suggesting that it folds in a "serpentine" manner into three extracellular and two intracellular loops. STRAP-1 mRNA and protein expression is maintained at high levels and throughout all stages of prostate cancer. STRAP-1 mRNA and/or protein is also overexpressed in certain other cancers, including bladder, colon, pancreatic and ovarian cancer. The function of the STRAP proteins is not known. They may be ion channels (from the presence of six transmembrane domains, a feature which is shared by certain ion channels) or gap-junction proteins (from immunohistochemical staining). STRAP-1 and STRAP-2 are cell-surface tumour antigens. Immunisation with a STRAP protein induces cellular and humoral immune responses against STRAP-expressing cells. STRAP proteins may be used to identify specific-binding agents, to produce anticancer vaccines and to generate specific antibodies. The antibodies may be used for detection, prognosis, and monitoring of cancers (or susceptibility to cancer), as therapeutic inhibitors or to target therapeutic agents to their site of action. STRAP nucleic acids may be used for recombinant protein production, as diagnostic and prognostic reagents, for identifying STRAP-expressing cells for screening inhibitors of STRAP expression and for therapeutic modulation/inhibition of STRAP expression. Since high levels of STRAP proteins are exposed on the cell surface, they are easily targeted by systemically administered agents, and because they are expressed mainly on prostatic epithelial cells, agents targeted to them should have minimal side effects on other tissues

Sequence 34 AA;

```

XX WPI; 2001-367804/38.
XX
XX New STEAP (six transmembrane epithelial antigen of the prostate)
XX PT proteins, expressed in human cancers, useful for detecting and treating
XX PT cancer.
XX
XX Example 19; Page 102; 187pp; English.
XX
XX The present invention relates to human six transmembrane epithelial
XX CC antigen of the prostate (STEAP) protein. STEAP is a member of cell
XX CC surface serpentine transmembrane antigens. STEAP gene is used in gene
XX CC therapy. Inhibiting the development or progression of a cancer (eg.
XX CC prostate, colon, bladder, lung, ovarian and pancreatic) expressing STEAP
XX CC or inhibiting growth or killing cells expressing STEAP in a patient,
XX CC comprises administering a vaccine composition to the patient. Treating a
XX CC patient with a cancer that expresses STEAP, or inhibiting growth or
XX CC killing cells expressing STEAP, comprises administering to the patient a
XX CC vector encoding single chain monoclonal antibody that comprises the
XX CC variable domains of the heavy and light chains of the monoclonal antibody
XX CC that specifically binds to STEAP, such that the vector delivers the
XX CC single chain monoclonal antibody coding sequence to the cancer cells and
XX CC the encoded single chain monoclonal antibody is expressed
XX CC intracellularly. The present sequence is extracellular loop of STEAP-1
XX CC suitable for cloning into pFC, which is used in the invention. STEAP-1
XX CC gene is located on chromosome 7p22.3
XX
XX Sequence 34 AA;
XX
XX Query Match 100.0%; Score 192; DB 4; Length 34;
XX Best Local Similarity 100.0%; Pred. No. 3.8e-19;
XX Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 34
XX | | | | | | | | | | | | | | | | | | | |
XX Db 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 34
XX
XX RESULT 3
XX AAB75315
XX ID AAB75315 standard; protein; 129 AA.
XX
XX AC AAB75315;
XX
XX DT 03-APR-2001 (first entry)
XX
XX DE Human secreted protein sequence encoded by gene 20 SEQ ID NO:134.
XX
XX KW Human; immunosuppressive; antiarthritic; antiarthritis; vasotropic; neuroproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
XX KW neuroprotective; antibacterial; virucide; fungicide; ophthalmological;
XX KW vulnerary; autoimmune disease; hyperproliferative disorder; cancer;
XX KW cardiovascular disorder; cerebrovascular disorder; infection;
XX KW nervous system disorder; ocular disorder; chemotaxis; food additive;
XX KW secreted protein.
XX
XX OS Homo sapiens.
XX
XX PN WC200077021-A1.
XX
XX PD 21-DEC-2000.
XX
XX PF 01-JUN-2000; 2000WO-US015135.
XX
XX PR 11-JUN-1999; 99US-0138632P.
XX
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PA (ROSE/) ROSEN C A.
XX
XX PI Rosen CA, Ruben SM, Komatsoulis GA;
XX
XX WPI; 2001-071257/08.
XX DR N-PSDB; AAF63808.
XX

```

```
XX PT Nucleic acid molecules encoding human secreted proteins, used in
XX PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
XX PT Parkinson's diseases and cancers.
XX PS Disclosure; Page 53-54; 530pp; English.
XX CC This invention relates to polynucleotide sequences AAF63789 - AAF63836
XX CC which encode human secreted proteins AAB75260 - AAB75287. Included in the
XX CC invention are protein sequences AAB75288 - AAB75341 which are fragments
XX CC of the secreted proteins and amino acid sequences with which these
XX CC fragments share homology. Examples of the activities of the proteins and
XX CC polynucleotides and the activities of their agonists and antagonists
XX CC include, immunosuppressive; antiarthritic; antirheumatic;
XX CC antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
XX CC neurotropic; neuroprotective; antibacterial; virucide; fungicide;
XX CC ophthalmological; and vulnary activity. The protein and polynucleotide
XX CC sequences, their agonists and antagonists may be useful for treating,
XX CC preventing and diagnosing diseases and disorders such as autoimmune
XX CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. cardiac
XX CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
XX CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
XX CC nervous system disorders e.g. Alzheimer's disease, infections caused by
XX CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection.
XX CC The polypeptides can also be used to aid wound healing and epithelial
XX CC cell proliferation, to prevent skin aging due to sunburn, to maintain
XX CC organs before transplantation, for supporting cell culture of primary
XX CC tissues, to regenerate tissues and in chemotaxis. The polypeptides can
XX CC also be used as a food additive or preservative to increase or decrease
XX CC storage capabilities. Included in the invention are oligonucleotides
XX CC AAF63780 - AAF63788 and peptide AAB75239 which are used in the
XX CC identification and characterisation of the DNA and protein sequences of
XX CC the invention
XX SQ Sequence 129 AA;
Query Match 100.0%; Score 192; DB 4; Length 129;
Best Local Similarity 100.0%; Pred. No. 1.8e-18;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 34
Db 60 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 93
RESULT 4
AAB75314
XX ID AAB75314 standard; protein; 129 AA.
XX AC AAB75314;
XX DT 03-APR-2001 (first entry)
XX DE Gene 20 human secreted protein homologous amino acid sequence #133.
XX KW Human; immunosuppressive; antiarthritic; antirheumatic; neurotropic;
XX KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
XX KW neuroprotective; antibacterial; virucide; fungicide; ophthalmological;
XX KW vulnary; autoimmune disease; hyperproliferative disorder; cancer;
XX KW cardiovascular disorder; cerebrovascular disorder; infection;
XX KW nervous system disorder; ocular disorder; chemotaxis; food additive;
XX KW secreted protein.
XX OS Homo sapiens.
XX PN WO200077021-A1.
XX PD 21-DEC-2000.
XX PF 01-JUN-2000; 2000WO-US015135.
XX PR 11-JUN-1999; 99US-0138632P.
XX XX
```

```
PA (HVA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
XX PI Rosen CA, Ruben SM, Komatsoulis GA;
XX DR WPI; 2001-071257/08.
XX PT Nucleic acid molecules encoding human secreted proteins, used in
XX PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
XX PT Parkinson's diseases and cancers.
XX PS Disclosure; Page 53-54; 530pp; English.
XX CC This invention relates to polynucleotide sequences AAF63789 - AAF63836
XX CC which encode human secreted proteins AAB75260 - AAB75287. Included in the
XX CC invention are protein sequences AAB75288 - AAB75341 which are fragments
XX CC of the secreted proteins and amino acid sequences with which these
XX CC fragments share homology. Examples of the activities of the proteins and
XX CC polynucleotides and the activities of their agonists and antagonists
XX CC include, immunosuppressive; antiarthritic; antirheumatic;
XX CC antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
XX CC neurotropic; neuroprotective; antibacterial; virucide; fungicide;
XX CC ophthalmological; and vulnary activity. The protein and polynucleotide
XX CC sequences, their agonists and antagonists may be useful for treating,
XX CC preventing and diagnosing diseases and disorders such as autoimmune
XX CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. cardiac
XX CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
XX CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
XX CC nervous system disorders e.g. Alzheimer's disease, infections caused by
XX CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection.
XX CC The polypeptides can also be used to aid wound healing and epithelial
XX CC cell proliferation, to prevent skin aging due to sunburn, to maintain
XX CC organs before transplantation, for supporting cell culture of primary
XX CC tissues, to regenerate tissues and in chemotaxis. The polypeptides can
XX CC also be used as a food additive or preservative to increase or decrease
XX CC storage capabilities. Included in the invention are oligonucleotides
XX CC AAF63780 - AAF63788 and peptide AAB75239 which are used in the
XX CC identification and characterisation of the DNA and protein sequences of
XX CC the invention
XX SQ Sequence 129 AA;
Query Match 100.0%; Score 192; DB 4; Length 129;
Best Local Similarity 100.0%; Pred. No. 1.8e-18;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 34
Db 60 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 93
RESULT 5
AAB75426
XX ID AAB75426 standard; protein; 254 AA.
XX AC AAB75426;
XX DT 31-JUL-2003 (first entry)
XX DE STEAP-1 variant 8P1D4 v.2 #2.
XX KW STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
XX KW cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
XX KW vaccine.
XX OS Homo sapiens.
XX PN WO2003022995-A2.
XX PD 20-MAR-2003.
XX PF 06-SEP-2002; 2002WO-US028371.
XX XX
```

PR 06-SEP-2001; 2001US-0317840P.
PR 05-APR-2002; 2002US-0370387P.
XX (AGEN-) AGENSYS INC.
XX Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;
XX WPI; 2003-313240/30.
DR
XX
XX New composition comprising a substance that modulates the status of a
PT STEAP-1-related protein, useful for treating and detecting cancer.
XX
XX Example 53; Page 169-170; 248pp; English.
XX
XX The invention describes a composition comprising a substance that
CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
CC any of the 15 sequences of 259 amino acids, given in the specification,
CC or a molecule that is modulated by the protein, where the status of the
CC cell that expresses the protein is modulated. The compositions, proteins,
CC polynucleotides and methods are useful for treating and detecting cancer.
CC The STEAP-1-related proteins are useful for generating cancer vaccines.
CC The polynucleotides are useful as tools for delineating, with greater
CC precision, cytogenetic abnormalities in the chromosomal region that
CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
CC the amino acid sequence of a variant of human six transmembrane
CC epithelial antigen of the prostate or STEAP-1
XX
SQ Sequence 254 AA;
Query Match 100.0%; Score 192; DB 6; Length 254;
Best Local Similarity 100.0%; Pred. No. 3.8e-18;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RRSYRYKLLNWAYQQVQNKEDAWIEHDVWRMEI 34
Db 185 RRSYRYKLLNWAYQQVQNKEDAWIEHDVWRMEI 218
RESULT 6
ABU98425
ID ABU98425 standard; protein; 254 AA.
XX
AC ABU98425;
XX
XX 31-JUL-2003 (first entry)
XX
XX STEAP-1 variant 8P1D4 v.1 #1.
XX
XX STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
KW cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
KW vaccine.
XX
XX Homo sapiens.
OS
XX WO2003022995-A2.
XX
XX 20-MAR-2003.
XX
XX 06-SEP-2002; 2002WO-US028371.
XX
XX 06-SEP-2001; 2001US-0317840P.
PR 05-APR-2002; 2002US-0370387P.
XX
XX (AGEN-) AGENSYS INC.
FA
XX Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;
XX WPI; 2003-313240/30.
DR
XX
XX New composition comprising a substance that modulates the status of a
PT STEAP-1-related protein, useful for treating and detecting cancer.
XX
XX Example 53; Page 169-170; 248pp; English.

XX
CC The invention describes a composition comprising a substance that
CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
CC any of the 15 sequences of 259 amino acids, given in the specification,
CC or a molecule that is modulated by the protein, where the status of the
CC cell that expresses the protein is modulated. The compositions, proteins,
CC polynucleotides and methods are useful for treating and detecting cancer.
CC The STEAP-1-related proteins are useful for generating cancer vaccines.
CC The polynucleotides are useful as tools for delineating, with greater
CC precision, cytogenetic abnormalities in the chromosomal region that
CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
CC the amino acid sequence of a variant of human six transmembrane
CC epithelial antigen of the prostate or STEAP-1
XX
SQ Sequence 254 AA;
Query Match 100.0%; Score 192; DB 6; Length 254;
Best Local Similarity 100.0%; Pred. No. 3.8e-18;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RRSYRYKLLNWAYQQVQNKEDAWIEHDVWRMEI 34
Db 185 RRSYRYKLLNWAYQQVQNKEDAWIEHDVWRMEI 218
RESULT 7
ABU98429
ID ABU98429 standard; protein; 255 AA.
XX
AC ABU98429;
XX
XX 31-JUL-2003 (first entry)
XX
XX STEAP-1 variant 8P1D4 v.3 #2.
DE
XX STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
KW cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
KW vaccine.
XX
XX Homo sapiens.
OS
XX WO2003022995-A2.
XX
XX 20-MAR-2003.
XX
XX 06-SEP-2002; 2002WO-US028371.
XX
XX 06-SEP-2001; 2001US-0317840P.
PR 05-APR-2002; 2002US-0370387P.
XX
XX (AGEN-) AGENSYS INC.
FA
XX Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;
XX WPI; 2003-313240/30.
DR
XX
XX New composition comprising a substance that modulates the status of a
PT STEAP-1-related protein, useful for treating and detecting cancer.
XX
XX Example 53; Page 172; 248pp; English.
XX
XX The invention describes a composition comprising a substance that
CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
CC any of the 15 sequences of 259 amino acids, given in the specification,
CC or a molecule that is modulated by the protein, where the status of the
CC cell that expresses the protein is modulated. The compositions, proteins,
CC polynucleotides and methods are useful for treating and detecting cancer.
CC The STEAP-1-related proteins are useful for generating cancer vaccines.
CC The polynucleotides are useful as tools for delineating, with greater
CC precision, cytogenetic abnormalities in the chromosomal region that
CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
CC the amino acid sequence of a variant of human six transmembrane
CC epithelial antigen of the prostate or STEAP-1
XX

```

; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685,166A
; CURRENT FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-685-166A-879

Query Match      100.0%; Score 86; DB 4; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WKMKPRNLEEDDY 15
Db      14 WKMKPRNLEEDDY 28

RESULT 3
US-09-134-000C-6624
; Sequence 6624, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6624
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-6624

Query Match      53.5%; Score 46; DB 4; Length 216;
Best Local Similarity 61.5%; Pred. No. 4;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      3 MKPRNLEEDDY 15
Db      125 IKPRYLEEGDY 137

RESULT 4
US-09-489-039A-11367
; Sequence 11367, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747

```

```

; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11367
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11367

Query Match      50.0%; Score 43; DB 4; Length 278;
Best Local Similarity 53.8%; Pred. No. 17;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      2 KMKPRNLEEDDY 14
Db      66 RLAPRRELEGDY 78

RESULT 5
US-09-134-001C-5278
; Sequence 5278, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5278
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5278

Query Match      49.4%; Score 42.5; DB 4; Length 256;
Best Local Similarity 64.3%; Pred. No. 19;
Matches 9; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

QY      2 KMKPRNLEEDDY 15
Db      234 KVKPRNN-KEDNYL 246

RESULT 6
US-09-485-632B-6
; Sequence 6, Application US/09485632B
; Patent No. 6605280
; GENERAL INFORMATION:
; APPLICANT: No. 6605280ick, Daniela
; APPLICANT: Dinarello, Charles
; APPLICANT: Rubinstein, Menachem
; APPLICANT: Kim, Soo Hyun
; TITLE OF INVENTION: Interleukin-18 Binding Proteins, their Preparation and
; TITLE OF INVENTION: Use
; FILE REFERENCE: 20993-001
; CURRENT APPLICATION NUMBER: US/09/485,632B
; CURRENT FILING DATE: 2000-10-12
; PRIOR APPLICATION NUMBER: IL98/00379
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: 125463
; PRIOR FILING DATE: 1998-07-22
; PRIOR APPLICATION NUMBER: 122134
; PRIOR FILING DATE: 1997-11-06
; PRIOR APPLICATION NUMBER: 121869
; PRIOR FILING DATE: 1997-09-29
; PRIOR APPLICATION NUMBER: 121639
; PRIOR FILING DATE: 1997-08-27
; PRIOR APPLICATION NUMBER: 121554

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 14:04:08 ; Search time 9.24419 Seconds
(without alignments)
83.770 Million cell updates/sec

Title: US-10-010-667A-2_COPY_14_28

Perfect score: 86

Sequence: 1 WKMKPRNLEDDYL 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PTUS COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	86	100.0	339	4	US-09-323-873A-2
2	86	100.0	339	4	US-09-685-166A-879
3	46	53.5	216	4	US-09-134-000C-6624
4	43	50.0	278	4	US-09-489-039A-11367
5	42.5	49.4	256	4	US-09-134-001C-5278
6	40	46.5	197	4	US-09-485-632B-6
7	40	46.5	221	3	US-09-247-373B-54
8	40	46.5	573	4	US-09-252-991A-31334
9	38	44.2	103	4	US-09-621-976-5168
10	38	44.2	193	4	US-09-582-379-4
11	38	44.2	345	1	US-08-403-866-2
12	38	44.2	464	4	US-09-252-991A-29449
13	38	44.2	540	4	US-09-134-000C-4765
14	38	44.2	555	4	US-09-107-532A-6422
15	38	44.2	968	3	US-08-651-999A-7
16	38	44.2	968	3	US-09-385-752-7
17	38	44.2	3969	3	US-08-061-376-5
18	37	43.0	25	1	US-08-468-709B-16
19	37	43.0	25	2	US-08-241-664B-16
20	37	43.0	72	4	US-09-543-681A-7461
21	37	43.0	79	4	US-09-621-976-6283
22	37	43.0	220	4	US-09-489-039A-13425
23	37	43.0	323	2	US-08-969-106-4
24	37	43.0	323	4	US-09-338-125-4
25	37	43.0	416	1	US-08-252-995D-2
26	37	43.0	416	2	US-08-834-108-2
27	37	43.0	464	1	US-08-252-995D-6

28 37 43.0 464 2 US-08-834-108-6 Sequence 6, Appli
29 37 43.0 522 4 US-09-446-301A-49 Sequence 49, Appli
30 37 43.0 522 4 US-09-089-932-49 Sequence 49, Appli
31 37 43.0 577 1 US-08-484-105-24 Sequence 24, Appli
32 37 43.0 577 1 US-08-484-106-24 Sequence 24, Appli
33 37 43.0 577 2 US-08-756-317-13 Sequence 13, Appli
34 37 43.0 605 4 US-09-394-645-2 Sequence 2, Appli
35 37 43.0 605 4 US-09-243-560B-2 Sequence 2, Appli
36 37 43.0 711 3 US-08-861-083-82 Sequence 82, Appli
37 37 43.0 711 4 US-09-536-924-82 Sequence 82, Appli
38 37 43.0 751 4 US-09-402-929-2 Sequence 2, Appli
39 37 43.0 752 4 US-09-402-929-2 Sequence 2, Appli
40 37 43.0 925 1 US-08-252-995D-4 Sequence 6, Appli
41 37 43.0 925 2 US-08-834-108-4 Sequence 4, Appli
42 37 43.0 1228 4 US-09-439-313-537 Sequence 537, App
43 37 43.0 1228 4 US-09-636-215-537 Sequence 537, App
44 37 43.0 1228 4 US-09-685-166A-537 Sequence 537, App
45 36 41.9 130 4 US-09-489-039A-7711 Sequence 7711, Ap

ALIGNMENTS

RESULT 1
US-09-323-873A-2
; Sequence 2, Application US/09323873A
; Patent No. 6329503
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Kahan Leong
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Steve Chappell Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; FILE REFERENCE: 129.16USU2
; CURRENT APPLICATION NUMBER: US/09/323,873A
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-323-873A-2

Query Match 100.0%; Score 86; DB 4; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WKMKPRNLEDDYL 15
Db 14 WKMKPRNLEDDYL 28

RESULT 2
US-09-685-166A-879
; Sequence 879, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.

Db 14 WKMKPRNLEDDYL 28
|||||

Search completed: July 12, 2004, 14:27:48
Job time : 26.5116 secs

```
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-010-667A-2
```

```
Query Match      100.0%; Score 86; DB 14; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 WKMKPRRNLEDDYL 15
        |||||
Db      14 WKMKPRRNLEDDYL 28
```

RESULT 13

```
US-10-205-823-397
; Sequence 397, Application US/10205823
; Publication No. US20030108963A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Monsey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; PRIOR FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 397
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-205-823-397
```

```
Query Match      100.0%; Score 86; DB 14; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 WKMKPRRNLEDDYL 15
        |||||
Db      14 WKMKPRRNLEDDYL 28
```

RESULT 14

```
US-10-144-678A-879
; Sequence 879, Application US/10144678A
; Publication No. US20030157089A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aljun
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Hepler, William T.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals Y de Bassols, Carlota
; APPLICANT: Foy, Teresa M.
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Deng, Ta
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C28
; CURRENT APPLICATION NUMBER: US/10/144,678A
; CURRENT FILING DATE: 2002-08-12
; NUMBER OF SEQ ID NOS: 1033
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-678A-879
```

```
Query Match      100.0%; Score 86; DB 14; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 WKMKPRRNLEDDYL 15
        |||||
Db      14 WKMKPRRNLEDDYL 28
```

RESULT 15

```
US-10-294-025-879
; Sequence 879, Application US/10294025
; Publication No. US20030185830A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Stolk, John A.
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C29
; CURRENT APPLICATION NUMBER: US/10/294,025
; CURRENT FILING DATE: 2002-11-12
; NUMBER OF SEQ ID NOS: 1038
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-294-025-879
```

```
Query Match      100.0%; Score 86; DB 14; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 WKMKPRRNLEDDYL 15
```



```
RESULT 9
US-10-408-009-2
; Sequence 2, Application US/10408009
; Publication No. US20040072196A1
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Stephen C. Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; FILE REFERENCE: 511582001603
; CURRENT APPLICATION NUMBER: US/10/408,009
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 09/455,486
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 09/323,873
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-009-2

Query Match      100.0%; Score 86; DB 12; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WKKMPRRNLEDDYL 15
DB      14 WKKMPRRNLEDDYL 28
      |||||
RESULT 10
US-10-012-896-879
; Sequence 879, Application US/10012896
; Publication No. US2002018255A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; APPLICANT: Wantanabe, Yoshihiro
; APPLICANT: Meagher, Madeleine Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C27
; CURRENT APPLICATION NUMBER: US/10/012,896
; CURRENT FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 1011
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 879

Query Match      100.0%; Score 86; DB 12; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WKKMPRRNLEDDYL 15
DB      14 WKKMPRRNLEDDYL 28
      |||||
RESULT 11
US-10-011-095-2
; Sequence 2, Application US/10011095
; Publication No. US20030045682A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Hubert, Rene S.
; APPLICANT: Leong, Kahan
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Saffran, Douglas C.
; APPLICANT: Mitchell, Steve Chappell
; TITLE OF INVENTION: ANTIBODIES IMMUNOSPECIFIC FOR STEAP1 (AS AMENDED)
; FILE REFERENCE: 511582001610
; CURRENT APPLICATION NUMBER: US/10/011,095
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: 09/323,873
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DNA
US-10-011-095-2

Query Match      100.0%; Score 86; DB 14; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WKKMPRRNLEDDYL 15
DB      14 WKKMPRRNLEDDYL 28
      |||||
RESULT 12
US-10-010-667A-2
; Sequence 2, Application US/10010667A
; Publication No. US20030055217A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Hubert, Rene S.
; APPLICANT: Leong, Kahan
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Saffran, Douglas C.
; APPLICANT: Mitchell, Steve Chappell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USES THEREOF
; FILE REFERENCE: 511582001601
; CURRENT APPLICATION NUMBER: US/10/010,667A
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: 09/323,873
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/087,520
```

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020187472A1 g6572948
US-09-802-520-11
```

```
Query Match      100.0%; Score 86; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 WKMKPRRNLEDDYL 15
        |||
Db      14 WKMKPRRNLEDDYL 28
```

RESULT 6

```
US-09-895-793-879
; Sequence 879, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C26
; CURRENT APPLICATION NUMBER: US/09/895,814
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-895-814-879
```

```
Query Match      100.0%; Score 86; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 WKMKPRRNLEDDYL 15
        |||
Db      14 WKMKPRRNLEDDYL 28
```

RESULT 8

```
US-10-205-267-13
; Sequence 13, Application US/10205267
; Publication No. US20030064397A1
; GENERAL INFORMATION:
; APPLICANT: Spaccake, Kimberly M.
; APPLICANT: Rickert Paula K.
; APPLICANT: Lal, Preeti G.
; APPLICANT: Ison, Craig H.
; TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN PROSTATE AND
; FILE REFERENCE: PV-0008 CIP
; CURRENT APPLICATION NUMBER: US/10/205,267
; CURRENT FILING DATE: 2002-07-24
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PERL Program
; SEQ ID NO 13
; LENGTH: 339
; TYPE: PRT
; ORGANISM:
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: Genbank ID No: g6572948
US-10-205-267-13
```

```
Query Match      100.0%; Score 86; DB 12; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 WKMKPRRNLEDDYL 15
        |||
Db      14 WKMKPRRNLEDDYL 28
```

RESULT 7

```
US-09-895-814-879
; Sequence 879, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
```

```
Query Match      100.0%; Score 86; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 WKMKPRRNLEDDYL 15
        |||
Db      14 WKMKPRRNLEDDYL 28
```

```
Query Match      100.0%; Score 86; DB 14; Length 15;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WKKKPRRNLEDDYL 15
Db 1 WKKKPRRNLEDDYL 15

RESULT 2
US-09-759-143-879
; Sequence 879, Application US/09759143
; Patent No. US2002002248A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C23
; CURRENT APPLICATION NUMBER: US/09/759,143
; CURRENT FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-759-143-879

Query Match      100.0%; Score 86; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WKKKPRRNLEDDYL 15
Db 14 WKKKPRRNLEDDYL 28

RESULT 3
US-09-780-669-879
; Sequence 879, Application US/09780669
; Patent No. US2002005197A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun

Query Match      100.0%; Score 86; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WKKKPRRNLEDDYL 15
Db 14 WKKKPRRNLEDDYL 28

RESULT 4
US-09-822-827-879
; Sequence 879, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C1
; CURRENT APPLICATION NUMBER: US/09/822,827
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-822-827-879

Query Match      100.0%; Score 86; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WKKKPRRNLEDDYL 15
Db 14 WKKKPRRNLEDDYL 28

RESULT 5
US-09-802-520-11
; Sequence 11, Application US/09802520
; Publication No. US20020187472A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Faris, Mary
; APPLICANT: Chen, Huel-Mei
; APPLICANT: Ison, Craig H.
; TITLE OF INVENTION: STEAP-RELATED PROTEIN
; FILE REFERENCE: PC-0037 US
; CURRENT APPLICATION NUMBER: US/09/802,520
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PERL Program
; SEQ ID NO 11
; LENGTH: 339
; TYPE: PRT
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 14:10:33 ; Search time 26.5116 Seconds
(without alignments)
176.480 Million cell updates/sec

Title: US-10-010-667a-2_COPY_14_28

Perfect score: 86

Sequence: 1 WWMKPRNLEDDYL 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1279676 seqs, 311918243 residues

Total number of hits satisfying chosen parameters: 1279676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA.*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	86	100.0	15	14	US-10-165-044-22
2	86	100.0	339	9	US-09-759-143-879
3	86	100.0	339	9	US-09-780-669-879
4	86	100.0	339	9	US-09-822-827-879
5	86	100.0	339	9	US-09-802-520-11
6	86	100.0	339	9	US-09-895-793-879
7	86	100.0	339	9	US-09-895-814-879
8	86	100.0	339	12	US-10-205-267-13
9	86	100.0	339	12	US-10-408-009-2
10	86	100.0	339	13	US-10-012-896-879
11	86	100.0	339	14	US-10-011-095-2
12	86	100.0	339	14	US-10-010-667A-2
13	86	100.0	339	14	US-10-205-823-397
14	86	100.0	339	14	US-10-144-678A-879
15	86	100.0	339	14	US-10-294-025-879

16	86	100.0	339	15	US-10-239-607-37	Sequence 37, Appl
17	86	100.0	339	15	US-10-295-027-714	Sequence 714, App
18	86	100.0	339	15	US-10-295-027-1347	Sequence 1347, Ap
19	86	100.0	368	12	US-10-425-114-72779	Sequence 72779, A
20	86	100.0	375	14	US-10-165-044-2	Sequence 2, Appli
21	46	53.5	561	12	US-10-424-599-248099	Sequence 248099,
22	45	52.3	2078	16	US-10-437-963-130585	Sequence 130585,
23	44	51.2	66	12	US-10-424-599-243298	Sequence 243298,
24	44	51.2	133	14	US-10-189-346-60	Sequence 60, Appl
25	43	50.0	137	16	US-10-437-963-128802	Sequence 128802,
26	43	50.0	398	16	US-10-437-963-165124	Sequence 165124,
27	43	50.0	2040	15	US-10-276-968-13	Sequence 13, Appl
28	42	48.8	411	16	US-10-437-963-204350	Sequence 204350,
29	42	48.8	442	15	US-10-369-493-17721	Sequence 17721, A
30	42	48.8	1018	9	US-09-801-574-32	Sequence 32, Appl
31	42	48.8	1019	9	US-09-801-574-76	Sequence 76, Appl
32	42	48.8	1019	14	US-10-132-861-2	Sequence 2, Appli
33	42	48.8	1137	12	US-10-425-114-63876	Sequence 63876, A
34	41.5	48.3	208	15	US-10-158-034-67	Sequence 67, Appl
35	41.5	48.3	305	12	US-10-425-114-43994	Sequence 43994, A
36	41.5	48.3	528	12	US-10-424-599-259292	Sequence 259292,
37	41.5	48.3	705	12	US-10-425-114-43788	Sequence 43788, A
38	41	47.7	32	12	US-10-671-403-203	Sequence 203, App
39	41	47.7	32	12	US-10-671-419-203	Sequence 203, App
40	41	47.7	32	12	US-10-670-844-203	Sequence 203, App
41	41	47.7	32	12	US-10-671-134-203	Sequence 203, App
42	41	47.7	32	12	US-10-673-098-203	Sequence 203, App
43	41	47.7	32	16	US-10-673-638-203	Sequence 203, App
44	41	47.7	32	16	US-10-673-127-203	Sequence 203, App
45	41	47.7	32	16	US-10-670-817-203	Sequence 203, App

ALIGNMENTS

RESULT 1

US-10-165-044-22
; Sequence 22, Application US/10165044
; Publication No. US20030149531A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Rene S. Hubert
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas Saffran
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Steven Chappell Mitchell
; APPLICANT: Mary Faris
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; FILE REFERENCE: 51158-20016.02
; CURRENT APPLICATION NUMBER: US/10/165,044
; CURRENT FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: US 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: US 60/091,183
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 09/323,873
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: US 09/455,486
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: WO 99/62941
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/33040
; PRIOR FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 15
; TYPE: PPT
; ORGANISM: Homo sapiens
US-10-165-044-22

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 14:01:53 ; Search time 7.5 seconds
(without alignments)
192.383 Million cell updates/sec

Title: US-10-010-667A-2_COPY_14_28

Perfect score: 86

Sequence: 1 WKMKPRNLEDDYL 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_78.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45.5	52.9	237	2 T50202	conserved hypothet
2	45	52.3	375	1 NRECD	ribonuclease III (
3	45	52.3	375	2 A99943	RNase D, processes
4	45	52.3	375	2 E85791	RNase D, processes
5	44	51.2	532	2 T02539	hypothetical prote
6	43	50.0	312	2 T15371	hypothetical prote
7	43	50.0	539	2 T39150	probable heat shoc
8	42	48.8	124	2 AH1754	bacteriophage prot
9	42	48.8	384	2 T49020	hypothetical prote
10	42	48.8	416	2 D96692	hypothetical prote
11	42	48.8	442	2 D36718	dihydroilpoamide S
12	42	48.8	451	2 T49021	hypothetical prote
13	42	48.8	847	2 J04836	alpha-glucuronidas
14	42	48.8	1832	2 AC2594	glutamate synthase
15	42	48.8	1858	2 D97376	hypothetical prote
16	41	47.7	260	2 S75010	ABC-type transport
17	41	47.7	271	2 C84550	hypothetical prote
18	41	47.7	403	2 G75287	NADH oxidase-relat
19	41	47.7	580	2 T21493	hypothetical prote
20	41	47.7	1187	2 T46637	transcription fact
21	40.5	47.1	639	2 F71258	probable heat shoc
22	40.5	47.1	658	2 S76909	hypothetical prote
23	40.5	47.1	1415	1 EDEBGA	immediate-early pr
24	40	46.5	36	2 B31872	retinoic acid-bind
25	40	46.5	138	2 I51265	xCRABP - African c
26	40	46.5	361	2 T49337	hypothetical prote
27	40	46.5	430	2 T23899	hypothetical prote
28	40	46.5	474	2 T00943	hypothetical prote
29	40	46.5	526	2 F85086	hypothetical prote

30	40	46.5	549	1 F69361	arginyl-tRNA synth
31	40	46.5	884	2 H83322	hypothetical prote
32	40	46.5	961	2 T01167	hypothetical prote
33	40	46.5	1268	2 T18955	hypothetical prote
34	39	45.3	90	2 H90665	probable Ogr famil
35	39	45.3	90	2 C85516	probable activator
36	39	45.3	371	2 C64499	isocitrate dehydro
37	39	45.3	426	2 E83981	pyruvate dehydrog
38	39	45.3	426	2 B95119	protein f286.2 [im
39	39	45.3	449	2 G89841	hypothetical prote
40	39	45.3	479	2 B69764	transcription regu
41	39	45.3	512	2 T23035	hypothetical prote
42	39	45.3	652	2 B84568	probable calmoduli
43	39	45.3	713	2 JCS870	poly(beta-D-mannur
44	39	45.3	976	2 A97104	zn-dependent metal
45	39	45.3	1019	2 T00117	dve protein - frul

ALIGNMENTS

RESULT 1

T50202

conserved hypothetical protein SPAC25B8.15c [imported] - fission yeast (Schizosaccharomyces pombe)
C/Species: Schizosaccharomyces pombe
C/Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jun-2000
C/Accession: T50202
R/Barrell, B.G.; Rajandream, M.A.; McDougall, R.C.; McLean, J.; Harris, D.
submitted to the EMBL Data Library, November 1999
A/Reference number: Z25045
A/Accession: T50202
A/Status: preliminary; translated from GR/EMBL/DBDJ
A/Molecule type: DNA
A/Residues: 1-237 <BAR>
A/Cross-references: EMBL:AL133225; PIDN:CAB61781.1; GSPDB:GN00066; SPDB:SPAC25B8.15c
A/Experimental source: strain 972h(-); cosmid c25B8
C/Genetics:
A/Gene: SPDB:SPAC25B8.15c
A/Map position: 1

Query Match 52.9%; Score 45.5; DB 2; Length 237;

Best Local Similarity 66.7%; Pred.No. 4.3;

Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

Qy 2 KMKPRNLEE-DDYL 15

Db 220 RMKPRKLRNMDYYL 234

RESULT 2

NRECD

ribonuclease III (EC 3.1.26.3) rnd - Escherichia coli (strain K-12)

N/Alternate names: ribonuclease D

C/Species: Escherichia coli

C/Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 01-Mar-2002

C/Accession: S01223; A30431; S41590; D64941; S42849; D21915

R/Zhang, J.; Deutscher, M.P.

Nucleic Acids Res. 16, 6265-6278, 1988

A/Title: Escherichia coli RNase D: sequencing of the rnd structural gene and purification

A/Reference number: S01223; MUID:88289400; PMID:3041371

A/Accession: S01223

A/Molecule type: DNA

A/Residues: 1-375 <ZHA>

A/Cross-references: EMBL:X07055; NID:g42770; PIDN:CAA30098.1; PID:g581215

A/Accession: A30431

A/Molecule type: protein

A/Residues: 1-6 <ZH2>

R/Fulda, M.; Heinz, E.; Wolter, F.P.

Mol. Gen. Genet. 242, 241-249, 1994

A/Title: The fadD gene of Escherichia coli K12 is located close to rnd at 39.6 min of t

A/Reference number: S41590; MUID:94150456; PMID:8107670

A/Accession: S41590

A/Molecule type: DNA

A:Residues: 1-38 <FUL>
A:Cross-references: GB:X70994; NID:G433478; PIDN:CAA50322.1; PID:G581071
A:Experimental source: strain K12
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cohen, A.; Rose, D.J.; Mau, B.; Shaoh, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: D64941
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-375 <BLAT>
A:Cross-references: GB:AE000274; GB:U00096; NID:G1788089; PIDN:AACT4874.1; PID:G1788105;
R:Fulda, M.
submitted to the EMBL Data Library, February 1993
A:Reference number: S42848
A:Accession: S42849
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 'L', 2-38 <FU2>
A:Cross-references: EMBL:X70994
C:Genetics:
A:Gene: rnd
A:Map position: 40 min
A:Start codon: TTG
C:Superfamily: ribonuclease D
C:Keywords: exonuclease; hydrolase
F:1-375/Product: ribonuclease D #status experimental <MAT>

Query Match 52.3%; Score 45; DB 1; Length 375;
Best Local Similarity 63.6%; Pred. No. 8.5;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WKMKPRNLEE 11
|||:|:|:|
Db 342 WKLKPNLNLE 352

RESULT 3
A99943
RNase D, processes tRNA precursor [imported] - Escherichia coli (strain O157:H7, substratum D, processes tRNA precursor [imported])
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C:Accession: A99943
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic reference number: A99629; MUID:21156231; PMID:11258796
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: A99943
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-375 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA035936.1; PID:G13361980; GSPDB:GN00154
A:Experimental source: strain O157:H7, substratum RMD 0509952
C:Genetics:
A:Gene: ECS2513
C:Superfamily: ribonuclease D

Query Match 52.3%; Score 45; DB 2; Length 375;
Best Local Similarity 63.6%; Pred. No. 8.5;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WKMKPRNLEE 11
|||:|:|:|
Db 342 WKLKPNLNLE 352

RESULT 4
B85791
RNase D, processes tRNA precursor [imported] - Escherichia coli (strain O157:H7, substratum D, processes tRNA precursor [imported])
C:Species: Escherichia coli

C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: E85791
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhiller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: E85791
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-375 <STO>
A:Cross-references: GB:AE005174; NID:G12515844; PIDN:AA056793.1; GSPDB:GN00145; UWGP:Z
A:Experimental source: strain O157:H7, substratum EDL933
C:Genetics:
A:Gene: rnd
C:Superfamily: ribonuclease D

Query Match 52.3%; Score 45; DB 2; Length 375;
Best Local Similarity 63.6%; Pred. No. 8.5;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WKMKPRNLEE 11
|||:|:|:|
Db 342 WKLKPNLNLE 352

RESULT 5
T02539
hypothetical protein At2g37730 [imported] - Arabidopsis thaliana
N:Alternate names: hypothetical protein F13M22.23
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 16-Feb-2001
C:Accession: T02539; D84796
R:Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, submitted to the EMBL Data Library, June 1998
A:Description: Arabidopsis thaliana chromosome II BAC F13M22 genomic sequence.
A:Reference number: Z14677
A:Accession: T02539
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-532 <ROU>
A:Cross-references: EMBL:AC004684; NID:G3236234; PID:G3236255
A:Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: D84796
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-532 <STO>
A:Cross-references: GB:AE002093; NID:G3236255; PIDN:AA023643.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g37730; F13M22.23
A:Map position: 2
A:Introns: 165/2; 295/3; 478/3
C:Superfamily: Arabidopsis hypothetical protein F13M22.23

Query Match 51.2%; Score 44; DB 2; Length 532;
Best Local Similarity 55.6%; Pred. No. 18;
Matches 10; Conservative 0; Mismatches 2; Indels 6; Gaps 1;

Qy 1 WKMKPRR-----NLEED 12
|||:|:|:|
Db 476 WKMAPRRQCCEIVNSEED 493

RESULT 6
T15371
hypothetical protein C01F1.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans

C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T15371
R:Johnson, D.
submitted to the EMBL Data Library, May 1996
A:Description: The sequence of C. elegans cosmid C01F1.
A:Reference number: Z18338
C:Accession: T15371
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-312 <OCH>
A:Cross-references: EMBL:U58761; NID:gl330391; PID:gl330396; PIDN:AAB00716.1; GSPDB:GN00
A:Experimental source: strain Bristol N2; clone C01F1
C:Genetics:
A:Gene: CESP:C01F1.2
A:Map position: 2
A:Introns: 36/2; 271/1

Query Match 50.0%; Score 43; DB 2; Length 312;
Best Local Similarity 63.6%; Pred. No. 15;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 PRNLEDDYL 15
|||:|||||:
Db 250 PRNLEDDYL 260

RESULT 7
T39150
probable heat shock transcription factor - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T39150
R:Oliver, K.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z21748
A:Accession: T39150
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-539 <OLI>
A:Cross-references: EMBL:Z99168; PIDN:CAB16301.1; GSPDB:GN000666; SPDB:SPAC8C9.14
A:Experimental source: strain 972H-; cosmid c8C9
C:Genetics:
A:Gene: SPDB:SPAC8C9.14
A:Map position: 1
A:Introns: 10/3; 40/3; 67/2; 86/3

Query Match 50.0%; Score 43; DB 2; Length 539;
Best Local Similarity 60.0%; Pred. No. 27;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 WKMKPRRNLEDDYL 15
|||:|||||:
Db 364 WKRPRIILLVDEL 378

RESULT 8
AH1754
bacteriophage protein homolog lin2581 [imported] - Listeria innocua (strain Clp11262)
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AH1754
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krefte, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AH1754
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-124 <GLA>

A:Cross-references: GB:AL592022; PIDN:CAC97808.1; PID:gl6415103; GSPDB:GN00178
A:Experimental source: strain Clp11262
C:Genetics:
A:Gene: lin2581

Query Match 48.8%; Score 42; DB 2; Length 124;
Best Local Similarity 46.7%; Pred. No. 8.6;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 WKMKPRRNLEDDYL 15
|||:|||||:
Db 15 WKRRKKILLRDEYL 29

RESULT 9
T49020
hypothetical protein F3C22.70 - Arabidopsis thaliana (fragment)
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C:Accession: T49020
R:Purnelle, B.; Masuy, D.; Goffeau, A.; Boutry, M.; Mewes, H.W.; Rudd, S.; Lemcke, K.; P
submitted to the Protein Sequence Database, April 2000
A:Reference number: Z25013
A:Accession: T49020
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-384 <PUR>
A:Cross-references: EMBL:AL353912; GSPDB:GN000061; ATSP:F3C22.70
A:Experimental source: cultivar Columbia; BAC clone F3C22
C:Genetics:
A:Gene: ATSP:F3C22.70
A:Map position: 3
A:Introns: 61/1; 252/3; 281/3

Query Match 48.8%; Score 42; DB 2; Length 384;
Best Local Similarity 42.9%; Pred. No. 28;
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 WKMKPRRNLEDDY 14
|||:|||||:
Db 46 WKLVNLEFDSDDY 59

RESULT 10
D96692
hypothetical protein T12I7.9 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: D96692
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, N.F.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.;
ansen, C.W.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D96692
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-416 <STO>
A:Cross-references: GB:AE005173; NID:gl1054573; PIDN:AAG27848.1; GSPDB:GN00141
C:Genetics:
A:Gene: T12I7.9
A:Map position: 1

Query Match 48.8%; Score 42; DB 2; Length 416;
Best Local Similarity 50.0%; Pred. No. 31;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 WKMKPRNLEEDDY 14
||| :|||
Db 40 WKYVPRDLDEADF 53

RESULT 11

D36718

dihydrolipoamide S-acetyltransferase (EC 2.3.1.12) precursor - Bacillus subtilis
N;Alternate names: pyruvate dehydrogenase complex, E2 component
C;Species: Bacillus subtilis

C;Date: 12-Apr-1991 #sequence_revision 12-Apr-1991 #text_change 20-Jun-2000
C;Accession: D36718, B69674

R;Hemilae; H.; Palva, A.; Paulin, L.; Arvidson, S.; Palva, I.

J. Bacteriol. 172, 5052-5063, 1990

A;Title: Secretory S complex of Bacillus subtilis: sequence analysis and identity to pyv

A;Reference number: A36718; MUID:90368558; PMID:1697575

A;Accession: D36718

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-442 <HEM>

A;Cross-references: GB:M57435; GB:M31542; NID:g143375; PIDN:AAA62683.1; PID:g143379

R;Kunat, P.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

C.; Bron, S.; Brouillet, S.; Bruch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch

A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Erttington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gall

isch, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.

Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,

A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel

Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,

A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron

teuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Togononi, A.; Tosato, V.; Uchiyama

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K

A;Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.

A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A;Reference number: A69580; MUID:98044033; PMID:9384377

A;Accession: B69674

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-442 <KUN>

A;Cross-references: GB:Z99111; GB:AL009126; NID:g2633699; PIDN:CAB13333.1; PID:g2633831

A;Experimental source: strain 168

C;Genetics:

A;Gene: pdhC

C;Superfamily: dihydrolipoamide acetyltransferase; lipoyl/biotin-binding homology

C;Keywords: acyltransferase; coenzyme A; oxidoreductase

F;4-77/Domain: lipoyl/biotin-binding homology <LBP>

F;413,417/Active site: His, Asp #status predicted

Query Match 48.8%; Score 42; DB 2; Length 442;
Best Local Similarity 46.7%; Pred. No. 33;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 WKMKPRNLEEDDY 15
||| :|||
Db 22 WYVPRNDEDDV 36

RESULT 12

T49021

hypothetical protein F3C22.80 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000

C;Accession: T49021

R;Purnelle, B.; Masuy, D.; Goffeau, A.; Boutry, M.; Mewes, H.W.; Rudd, S.; Lemcke, K.; M

submitted to the Protein Sequence Database, April 2000

A;Reference number: Z25013

A;Accession: T49021

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-451 <PUR>

A;Cross-references: EMBL:AL353912; GSPDB:GN00061; ATSP:F3C22.80

A;Experimental source: cultivar Columbia; BAC clone F3C22

C;Genetics:

A;Gene: ATSP:F3C22.80

A;Map position: 3

A;Introns: 4/2; 294/3; 339/3

Query Match 48.8%; Score 42; DB 2; Length 451;
Best Local Similarity 42.9%; Pred. No. 33;
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 WKMKPRNLEEDDY 14
||| :|||
Db 52 WKLVNLEPDDY 65

RESULT 13

JC4836

alpha-glucuronidase (EC 3.2.1.-) precursor - fungus (Trichoderma reesei)

N;Alternate names: GLRI

C;Species: Trichoderma reesei

C;Date: 15-Aug-1996 #sequence_revision 15-Oct-1996 #text_change 03-Nov-2003

C;Accession: JC4836

R;Margolles-Clark, E.; Saitohimo, M.; Slika-aho, M.; Penttilae, M.

Gene 172, 171-172, 1996

A;Title: The alpha-glucuronidase-encoding gene of Trichoderma reesei.

A;Reference number: JC4836; MUID:96257277; PMID:8654984

A;Accession: JC4836

A;Molecule type: mRNA

A;Residues: 1-847 <MAR>

A;Cross-references: EMBL:Z68706; NID:g1419337; PID:e218512; PID:g1419338

C;Comment: This enzyme releases glucuronic acid attached to xylose units of xylan.

C;Genetics:

A;Gene: glr1

C;Superfamily: alpha-glucuronidase

C;Keywords: glycosidase; hydrolase

F;1-19/Domain: signal sequence #status predicted <SIG>

F;20-847/Product: alpha-glucuronidase #status predicted <MAP>

Query Match 48.8%; Score 42; DB 2; Length 847;
Best Local Similarity 53.8%; Pred. No. 65;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 WKMKPRNLEEDDY 14
||| :|||
Db 108 KLSPLNLEEDDY 120

RESULT 14

AC2594

glutamate synthase large subunit gltB [imported] - Agrobacterium tumefaciens (strain C58)

C;Species: Agrobacterium tumefaciens

C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002

C;Accession: AC2594

R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo,

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell

; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

ster, E.W.

A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A;Reference number: AB2577; MUID:21608550; PMID:11743193

A;Accession: AC2594

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1832 <KUR>

A;Cross-references: GB:AE008688; PIDN:AAL41169.1; PID:g17738468; GSPDB:GN00186

A;Experimental source: strain C58 (Dupont)

C;Genetics:

A;Gene: gltB

A;Map position: circular chromosome

Query Match 48.8%; Score 42; DB 2; Length 1832;
Best Local Similarity 53.8%; Pred. No. 1.4e+02;

This Page Blank (uspto)

this Page Blank (uspto)

; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; FILE REFERENCE: CAPTURE AGENTS, AND USES THEREOF
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 14765
; LENGTH: 996
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-14765

Query Match 48.8%; Score 42; DB 6; Length 996;
Best Local Similarity 53.8%; Pred. No. 96;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 WKMKPRNLEEDD 13
Db 622 WKLVKSNLEED 634

Search completed: July 12, 2004, 14:25:04
Job time : 10.4186 secs


```
; CURRENT APPLICATION NUMBER: US/10/170,205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18326
; LENGTH: 349
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-18326
```

```
Query Match      100.0%; Score 86; DB 6; Length 349;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 WKMKPRRNLEDDYL 15
Db 24 WKMKPRRNLEDDYL 38
|||||
```

RESULT 7

```
US-10-857-785-35
; Sequence 35, Application US/10857785
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Stephen C. Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; FILE REFERENCE: 511582001607
; CURRENT APPLICATION NUMBER: US/10/857,785
; CURRENT FILING DATE: 2004-05-28
; PRIOR APPLICATION NUMBER: US 09/455,486
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: US 09/323,873
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: US 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: US 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-857-785-35
```

```
Query Match      87.8%; Score 75.5; DB 6; Length 16;
Best Local Similarity 93.8%; Pred. No. 2.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
```

```
QY 1 WKMKP-RRNLEDDYL 15
Db 1 WKMKPRRNLEDDYL 16
|||||
```

RESULT 8

```
US-10-856-109-35
; Sequence 35, Application US/10856109
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Stephen C. Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; FILE REFERENCE: 511582001606
; CURRENT APPLICATION NUMBER: US/10/856,109
; CURRENT FILING DATE: 2004-05-28
; PRIOR APPLICATION NUMBER: US 09/455,486
```

```
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: US 09/323,873
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: US 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: US 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-856-109-35
```

```
Query Match      87.8%; Score 75.5; DB 6; Length 16;
Best Local Similarity 93.8%; Pred. No. 2.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
```

```
QY 1 WKMKP-RRNLEDDYL 15
Db 1 WKMKPRRNLEDDYL 16
|||||
```

RESULT 9

```
US-10-752-421-35
; Sequence 35, Application US/10752421
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Stephen C. Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; FILE REFERENCE: 511582001612
; CURRENT APPLICATION NUMBER: US/10/752,421
; CURRENT FILING DATE: 2004-01-05
; PRIOR APPLICATION NUMBER: US 09/455,486
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: US 09/323,873
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: US 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: US 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-752-421-35
```

```
Query Match      87.8%; Score 75.5; DB 6; Length 16;
Best Local Similarity 93.8%; Pred. No. 2.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
```

```
QY 1 WKMKP-RRNLEDDYL 15
Db 1 WKMKPRRNLEDDYL 16
|||||
```

RESULT 10

```
US-10-724-972A-5260
; Sequence 5260, Application US/10724972A
; GENERAL INFORMATION:
; APPLICANT: Doucette-Stamm, Lynn
; APPLICANT: Bush, David
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PATH03-16
; CURRENT APPLICATION NUMBER: US/10/724,972A
```

```
; PRIOR APPLICATION NUMBER: PCT/US01/09410
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,929
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-239-607A-37

Query Match          100.0%; Score 86; DB 6; Length 339;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WKMKPRNLEDDYL 15
Db      14 WKMKPRNLEDDYL 28

RESULT 3
US-10-857-785-2
; Sequence 2, Application US/10857785
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Stephen C. Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USES THEREOF
; FILE REFERENCE: 511582001607
; CURRENT APPLICATION NUMBER: US/10/857,785
; CURRENT FILING DATE: 2004-05-28
; PRIOR APPLICATION NUMBER: US 09/455,486
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: US 09/323,873
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: US 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: US 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-857-785-2

Query Match          100.0%; Score 86; DB 6; Length 339;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WKMKPRNLEDDYL 15
Db      14 WKMKPRNLEDDYL 28

RESULT 4
US-10-856-109-2
; Sequence 2, Application US/10856109
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Stephen C. Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USES THEREOF
; FILE REFERENCE: 511582001606
; CURRENT APPLICATION NUMBER: US/10/856,109
```

```
; CURRENT FILING DATE: 2004-05-28
; PRIOR APPLICATION NUMBER: US 09/455,486
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: US 09/323,873
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: US 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: US 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-856-109-2

Query Match          100.0%; Score 86; DB 6; Length 339;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WKMKPRNLEDDYL 15
Db      14 WKMKPRNLEDDYL 28

RESULT 5
US-10-752-421-2
; Sequence 2, Application US/10752421
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Stephen C. Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USES THEREOF
; FILE REFERENCE: 511582001612
; CURRENT APPLICATION NUMBER: US/10/752,421
; CURRENT FILING DATE: 2004-01-05
; PRIOR APPLICATION NUMBER: US 09/455,486
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: US 09/323,873
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: US 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: US 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-752-421-2

Query Match          100.0%; Score 86; DB 6; Length 339;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WKMKPRNLEDDYL 15
Db      14 WKMKPRNLEDDYL 28

RESULT 6
US-10-170-205E-18326
; Sequence 18326, Application US/10170205E
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; TITLE OF INVENTION: CAPTURE AGENTS, AND USES THEREOF
; FILE REFERENCE: CL001381
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 14:08:08 ; Search time 9.4186 Seconds
(without alignments)
118.046 Million cell updates/sec

Title: US-10-010-667a-2_COPY_14_28

Perfect score: 86

Sequence: 1 WWMKPRNLEDDYL 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 393080 seqs, 74121830 residues

Total number of hits satisfying chosen parameters: 393080

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Pending Patents AA New.*
- 1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep.*
 - 2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
 - 3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
 - 4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
 - 5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
 - 6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep.*
 - 7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match Length	DB ID	Description
1	86	100.0	339	6	US-10-821-801-741
2	86	100.0	339	6	US-10-239-607A-37
3	86	100.0	339	6	US-10-857-785-2
4	86	100.0	339	6	US-10-856-109-2
5	86	100.0	339	6	US-10-752-421-2
6	86	100.0	349	6	US-10-170-205E-18326
7	75.5	87.8	16	6	US-10-857-785-35
8	75.5	87.8	16	6	US-10-856-109-35
9	75.5	87.8	16	6	US-10-752-421-35
10	42.5	49.4	256	6	US-10-724-972A-5260
11	42	48.8	331	1	PCT-US04-19153-25
12	42	48.8	331	6	US-10-812-849-25
13	42	48.8	817	6	US-10-767-701-4432
14	42	48.8	897	6	US-10-170-205E-14827
15	42	48.8	996	6	US-10-170-205E-14765
16	42	48.8	1832	7	US-60-581-351-5639
17	42	48.8	1858	7	US-60-581-351-5641
18	41	47.7	32	5	US-09-716-964B-203
19	41	47.7	32	6	US-10-671-106-203
20	41	47.7	327	6	US-10-170-205E-23733
21	40	46.5	72	5	US-09-248-796A-21475
22	40	46.5	197	5	US-09-790-338B-6
23	40	46.5	234	6	US-10-170-205E-16541
24	40	46.5	317	6	US-10-767-471-744
25	40	46.5	317	6	US-10-767-471-748
26	40	46.5	317	6	US-10-170-205E-28535

27	40	46.5	360	6	US-10-767-471-747	Sequence 747, App
28	40	46.5	360	6	US-10-170-205E-27412	Sequence 27412, A
29	40	46.5	365	7	US-60-556-841-5722	Sequence 5722, App
30	40	46.5	465	6	US-10-767-471-745	Sequence 745, App
31	40	46.5	770	6	US-10-767-471-750	Sequence 750, App
32	40	46.5	770	6	US-10-821-801-848	Sequence 848, App
33	40	46.5	1012	6	US-10-767-471-749	Sequence 749, App
34	40	46.5	1012	6	US-10-170-205E-27413	Sequence 27413, A
35	39.5	45.9	490	1	PCT-US03-24982A-125	Sequence 125, App
36	39.5	45.9	687	7	US-60-551-161-260	Sequence 260, App
37	39	45.3	142	7	US-60-565-632-11886	Sequence 11886, A
38	39	45.3	142	7	US-60-579-062-11886	Sequence 11886, A
39	39	45.3	321	6	US-10-170-205E-24904	Sequence 24904, A
40	39	45.3	349	7	US-60-546-745-23	Sequence 23, Appl
41	39	45.3	485	1	PCT-US04-19924-32	Sequence 32, Appl
42	39	45.3	485	6	US-10-873-467-32	Sequence 32, Appl
43	39	45.3	485	7	US-60-551-121-100	Sequence 100, App
44	39	45.3	485	7	US-60-581-351-13766	Sequence 13766, A
45	39	45.3	528	1	PCT-US02-22858A-357	Sequence 357, App

ALIGNMENTS

RESULT 1

US-10-821-801-741
; Sequence 741, Application US/10821801
; GENERAL INFORMATION:
; APPLICANT: Hinzmann, Bernd
; APPLICANT: Rosenthal, Andre
; APPLICANT: Hermann, Klaus
; APPLICANT: Heiden, Esmeralda
; APPLICANT: Pilarsky, Christian
; APPLICANT: Brummendorf, Thomas
; APPLICANT: Staub, Eike
; APPLICANT: Ropcke, Stefan
; APPLICANT: Mennerich, Detlev
; APPLICANT: Kinnemann, Henrik
; APPLICANT: Li, Xinzhong
; TITLE OF INVENTION: Human nucleic acid sequences from lung tumours
; FILE REFERENCE: 00154/002001
; CURRENT APPLICATION NUMBER: US/10/821,801
; CURRENT FILING DATE: 2004-04-09
; PRIOR APPLICATION NUMBER: DE 103 16 701.3
; PRIOR FILING DATE: 2003-04-09
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 741
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-821-801-741

Query Match 100.0%; Score 86; DB 6; Length 339;
Best Local Similarity 100.0%; Pred.No. 1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WWMKPRNLEDDYL 15
Db 14 WWMKPRNLEDDYL 28

RESULT 2

US-10-239-607A-37
; Sequence 37, Application US/10239607A
; GENERAL INFORMATION:
; APPLICANT: Saatcioglu, Fahri
; TITLE OF INVENTION: Novel Prostate-Specific or
; TITLE OF INVENTION: Testis-Specific Nucleic Acid Molecules, Polypeptides, and
; TITLE OF INVENTION: Diagnostic and Therapeutic Methods
; FILE REFERENCE: 50218/003002
; CURRENT APPLICATION NUMBER: US/10/239,607A
; CURRENT FILING DATE: 2003-03-20

This Page Blank (uspto)

This Page Blank (uspto)

Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 MKPRNLEDDYL 15
::|:|:|:|
Db 356 LRPLRTVETDDYL 368

RESULT 15

D97376
Hypothetical protein AGR_C_235 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
C:Accession: D97376
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorllo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: D97376
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1858 <KUR>
A:Cross-references: GB:AE007869; PIDN:AAK85965.1; PID:gl5155024; GSPDB:GN00169
C:Genetics:
A:Gene: AGR_C_235
A:Map position: circular chromosome

Query Match 48.8%; Score 42; DB 2; Length 1858;

Best Local Similarity 53.8%; Pred. No. 1.5e+02;

Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 MKPRNLEDDYL 15
::|:|:|:|
Db 382 LRPLRTVETDDYL 394

Search completed: July 12, 2004, 14:11:17
Job time : 8.5 secs

This Page Blank (uspto)

This Page Blank (uspto)

Search completed: July 12, 2004, 14:08:02
Job time : 5.7093 secs

```
KW Complete proteome.
FT SITE 113 123 "HIGH" REGION.
SQ SEQUENCE 549 AA; 62861 MW; 39DF41CC0B9AD210 CRC64;

Query Match 46.5%; Score 40; DB 1; Length 549;
Best Local Similarity 46.2%; Pred. No. 46;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 WKMKPRRNLEEDD 13
Db 377 WKILSERDMEDE 389

RESULT 15
GPT1_CANAL
ID GPT1_CANAL STANDARD; PRT; 553 AA.
AC O74248;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Putative polyamine transporter.
GN GPT1.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21179897; PubMed=11284011;
RA McNemar M.D.; Gorman J.A.; Buckley H.R.;
RT "Isolation of a gene encoding a putative polyamine transporter from
RT Candida albicans, GPT1."
RL Yeast 18:555-561(2001).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- SIMILARITY: Belongs to the amino acid permease family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF080132; AAC31569.1; -.
DR InterPro; IPR002293; AA/rel_permease1.
DR InterPro; IPR004840; AAC_permease.
DR InterPro; IPR004841; Permease_region.
DR Pfam; PF00324; aa_permeases; 1.
DR PROSITE; PS00218; AMINO_ACID_PERMEASE_1; FALSE_NEG.
KW Transport; Transmembrane.
FT TRANSMEM 43 63 POTENTIAL.
FT TRANSMEM 68 88 POTENTIAL.
FT TRANSMEM 133 153 POTENTIAL.
FT TRANSMEM 163 183 POTENTIAL.
FT TRANSMEM 190 210 POTENTIAL.
FT TRANSMEM 240 260 POTENTIAL.
FT TRANSMEM 280 300 POTENTIAL.
FT TRANSMEM 333 353 POTENTIAL.
FT TRANSMEM 397 417 POTENTIAL.
FT TRANSMEM 453 473 POTENTIAL.
FT TRANSMEM 485 505 POTENTIAL.
SQ SEQUENCE 553 AA; 60633 MW; 197DACB76C5FBC9E CRC64;

Query Match 46.5%; Score 40; DB 1; Length 553;
Best Local Similarity 46.2%; Pred. No. 46;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 WKMKPRRNLEEDD 13
Db 508 WYVGPRTNISEED 520
```

RT "Interleukin-18 binding protein: a novel modulator of the Th1 cytokine
response.";
RL Immunity 10:127-136(1999).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM A), SEQUENCE OF 29-33; 37-43; 46-52;
RS 55-58; 61-72; 89-93; 105-142 AND 149-167, AND TISSUE SPECIFICITY.
RX MEDLINE=99192308; PubMed=10094485;
RA Aizawa Y., Akita K., Tanitai M., Torigoe K., Mori T., Nishida Y.,
RA Ushio S., Nukada Y., Tanimoto T., Ikegami H., Ikeda M., Kurimoto M.;
RT "Cloning and expression of interleukin-18 binding protein.";
RL FEBS Lett. 445:338-342(1999).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM A).
RX MEDLINE=99263157; PubMed=10329540;
RA Xiang Y., Moss B.;
RT Identification of human and mouse homologs of the MC51L-53L-54L
family of secreted glycoproteins encoded by the Mollusca contagiosum
poxvirus.";
RL Virology 257:297-302(1999).
CC -!- FUNCTION: Binds to IL-18 and inhibits its activity. Functions as
an inhibitor of the early Th1 cytokine response.
CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=C; Synonyms=IL-18BPC;
CC IsoId=O95998-1; Sequence=Displayed;
CC Name=A; Synonyms=IL-18BPA;
CC IsoId=O95998-2; Sequence=VSP_002514;
CC Name=B; Synonyms=IL-18BPB;
CC IsoId=O95998-3; Sequence=VSP_002515, VSP_002516;
CC -!- TISSUE SPECIFICITY: Strongly expressed in heart, lung, placenta
and spleen.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; AF110798; AAD17187.1; -
CC EMBL; AF110798; AAD17188.1; -
CC EMBL; AF110798; AAD17189.1; -
CC EMBL; AF110799; AAD17190.1; -
CC EMBL; AF110800; AAD17191.1; -
CC EMBL; AF110801; AAD17192.1; -
CC EMBL; AF110804; BAA76374.1; -
CC EMBL; AF122906; AAD41051.1; ALT_INIT.
CC GenBank; HGNC:5987; IL18BP.
CC MIM; 604113; -
CC GO; GO:0005625; C:soluble fraction; TAS.
CC GO; GO:0042007; F:interleukin-18 binding; IDA.
CC GO; GO:0004871; F:signal transducer activity; TAS.
CC GO; GO:0006955; P:immune response; IDA.
CC InterPro; IPR007110; IG-like.
CC PROSITE; PS00835; IG_LIKE; 1.
CC Immunoglobulin domain; Glycoprotein; Signal; Alternative splicing.
KW SIGNAL 1 28
FT CHAIN 29 197 INTERLEUKIN-18 BINDING PROTEIN.
FT DOMAIN 63 164 IG-LIKE C2-TYPE.
FT DISULFID 84 148 BY SIMILARITY.
FT CARBOHYD 77 77 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 92 92 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 101 101 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 145 145 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARSPLIC 168 197 VSPRRGLQGEELCFHMGKGLCOSSL -> AGLRATL
PPTQEARLPSSHSPPQQG (in isoform A).
/FTid=VSP_002514.
FT FT 113 NGTSLSCVACSRRPFSILYMLNGSFIEHLGRILW ->
FT FT SWAEGNLAPHRSPALQPQQSTAGRLSLSTGPAARQ (in
isoform B).
FT FT

FT FT /FTid=VSP_002515.
FT Missing (in isoform B).
FT /FTid=VSP_002516.
SQ SEQUENCE 197 AA; 21698 MW; F1DE626AB2B285E1 CRC64;
Query Match 46.5%; Score 40; DB 1; Length 197;
Best Local Similarity 46.7%; Pred. No. 16;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY 1 WWMKPRRLNEEDDYLL 15
Db 167 WYRSPARGLQGEEL 181
RESULT 14
SYR_ARCFU
ID SYR_ARCFU STANDARD; PRT; 549 AA.
AC Q29368;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Arginyl-tRNA synthetase (EC 6.1.1.19) (Arginine--tRNA ligase) (ARGRS).
GN ARGS OR AF0894.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Kleck H.-P., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
CC -!- CATALYTIC ACTIVITY: ATP + L-arginine + tRNA(Arg) = AMP +
diphosphate + L-arginyl-tRNA(Arg).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; AE001042; AAB90346.1; -
CC EMBL; F69361; F69361.
CC TIGR; AF0894; -
CC HAMAP; MF_00123; -; 1.
CC InterPro; IPR001278; Arg_tRNA-synt_1c.
CC InterPro; IPR005148; N.
CC InterPro; IPR008909; tRNA-synt_1d_C.
CC InterPro; IPR001412; tRNA-synt_1.
CC Pfam; PF03485; N-Arg; 1.
CC Pfam; PF00750; tRNA-synt_id; 1.
CC Pfam; PF05746; tRNA-synt_id_C; 1.
CC PRINTS; PR01038; TRNASYNTHARG.
CC TIGRPFAM; TIGR00456; args; 1.
CC PROSITE; PS00178; AA_tRNA_LIGASE_I; FALSE NEG.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;

RESULT 12

```
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR008271; Ser Thr_pkin_AS.
DR InterPro; IPR002999; Tudor.
DR Pfam; PF00069; Pkinase; 1.
DR Pfam; PF00867; Tudor; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00333; Tudor; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; FALSE NEG.
DR PROSITE; PS00108; PROTEIN KINASE ST; FALSE_NEG.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS50304; Tudor; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Coiled coil.
FT DOMAIN 78 137 TUDOR.
FT DOMAIN 195 242 TNASE-LIKE.
FT DOMAIN 298 355 COILED COIL (POTENTIAL).
FT DOMAIN 710 1019 PROTEIN KINASE.
FT BIND 716 724 ATP (BY SIMILARITY).
FT BINDING 737 737 ATP (BY SIMILARITY).
FT CONFLICT 300 300 I -> F (IN REF. 2).
FT CONFLICT 509 509 F -> Y (IN REF. 2).
FT CONFLICT 621 623 KKI -> NKS (IN REF. 2).
FT CONFLICT 715 715 Y -> C (IN REF. 2).
FT CONFLICT 820 820 V -> A (IN REF. 2).
FT CONFLICT 948 948 K -> I (IN REF. 2).
FT CONFLICT 963 963 A -> G (IN REF. 2).
FT CONFLICT 1000 1000 T -> P (IN REF. 2).
FT CONFLICT 1010 1010 P -> T (IN REF. 2).
SQ SEQUENCE 1019 AA; 115729 MW; 5B58244A98D1590A CRC64;

Query Match 48.8%; Score 42; DB 1; Length 1019;
Best Local Similarity 53.8%; Pred. No. 41;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 WKMKPRNLEEDD 13
DB 645 WKLVKSNLEESD 657

RESULT 9
HTPG TREPA
ID HTPG_TREPA STANDARD; PRT; 639 AA.
AC O83949;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chaperone protein htpg (Heat shock protein htpg) (High temperature
DE Protein G)
DE HTPG OR TP0984.
OS Treponema pallidum.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nichols;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA McDonald L., Artiach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete."
RL Science 281:375-388(1998).
CC -!- FUNCTION: Molecular chaperone. Has ATPase activity (By
CC similarity).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the heat shock protein 90 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
-----
```

```
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; AE001265; AAC65938.1; -.
CC PIR; F71258; F71258.
CC HSSP; P02829; 1AH8.
CC TIGR; TP0984; -.
CC HAMAP; MF_00505; -. 1.
CC InterPro; IPR003594; ATPbind_ATPase.
CC InterPro; IPR001404; Hsp90.
CC Pfam; PF02518; HATPase_C; 1.
CC Pfam; PF00183; HSP90; 2.
CC PRINTS; PR00775; HEATSHOCK90.
CC SMART; SM00387; HATPase_C; 1.
CC PROSITE; PS00298; HSP90; 1.
KW Chaperone; ATP-binding; Heat shock; Complete proteome.
FT DOMAIN 1 348 A; SUBSTRATE-BINDING (BY SIMILARITY).
FT DOMAIN 349 565 B (BY SIMILARITY).
FT DOMAIN 566 639 C.
SQ SEQUENCE 639 AA; 72937 MW; 3E8FDBAC2282C31D CRC64;

Query Match 47.1%; Score 40.5; DB 1; Length 639;
Best Local Similarity 50.0%; Pred. No. 45;
Matches 7; Conservative 4; Mismatches 2; Indels 1; Gaps 1;

QY 1 WKMKPRNLEEDDY 14
DB 246 WK-RPKSELKEDY 258

RESULT 10
ICP4 HSVMG
ID ICP4_HSVMG STANDARD; PRT; 1415 AA.
AC Q02362;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Trans-acting transcriptional activator protein ICP4 (Immediate-early
DE protein IE175).
DE ICP4.
GN ICP4.
OS Marek's disease herpesvirus (strain GA) (MDHV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=10388;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92351564; PubMed=1322594;
RA Anderson A.S., Francesconi A., Morgan R.W.;
RT "Complete nucleotide sequence of the Marek's disease virus ICP4
RT gene."
RL Virology 189:657-667(1992).
CC -!- FUNCTION: THIS IE PROTEIN IS A MULTIFUNCTIONAL PROTEIN CAPABLE
CC OF MIGRATING TO THE NUCLEUS, BINDING TO DNA, TRANS-ACTIVATING
CC OTHER VIRAL GENES, AND AUTOREGULATING ITS OWN SYNTHESIS. IT IS
CC REQUIRED FOR THE SWITCH FROM IMMEDIATE-EARLY TO EARLY MODE OF
CC GENE EXPRESSION.
CC -!- SUBCELLULAR LOCATION: Nucleus of infected cells.
CC -!- PTM: A long stretch of serine residues may be a major site of
CC phosphorylation.
CC -!- SIMILARITY: BELONGS TO THE ICP4/IE140/IE180 FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
```

CC -!- SIMILARITY: BELONGS TO FAMILY 67 OF GLYCOSYL HYDROLASES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z68706; CAA92949.1; --
DR PIR; JC4836; JC4836.
DR InterPro; IPR005154; Glyco_hydro_67.
DR Pfam; PF03648; Glyco_hydro_67; 1.
KW Xylan degradation; Hydrolase; Glycosidase; Glycoprotein; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 847 ALPHA-GLUCURONIDASE.
FT CARBOHYD 52 52 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 238 238 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 321 321 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 353 353 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 586 586 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 692 692 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 740 740 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 767 767 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 847 AA; 93424 MW; 214AAE487382FDB4 CRC64;
Query Match 48.8%; Score 42; DB 1; Length 847;
Best Local Similarity 53.8%; Pred. No. 34;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 2 KMKPRNLEEDY 14
Db 108 KLSPLKLNKEDGY 120
[: : : : :]
[: : : : :]
RESULT 7
ST31_MOUSE
ID ST31_MOUSE STANDARD; PRT; 1018 AA.
AC Q99MW1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serine/threonine protein kinase 31 (EC 2.7.1.37).
GN STK31.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=21175748; PubMed=11279525;
RA Wang P.J., McCarrey J.R., Yang F., Page D.C.;
RT "An abundance of X-linked genes expressed in spermatogonia.";
RL Nat. Genet. 27:422-426(2001).
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- TISSUE SPECIFICITY: Testis specific. Expressed only in male germ
CC cells.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC -!- SIMILARITY: Contains 1 Tudor domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF285580; AAK31959.1; --
DR MGD; MG1:1924735; SPk31.
DR InterPro; IPR008191; Maternal_tudor.

DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002999; Tudor.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00567; Tudor; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00333; Tudor; 1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS0304; Tudor; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Coiled coil.
FT DOMAIN 78 137 TUDOR.
FT DOMAIN 105 245 TNASE-LIKE.
FT DOMAIN 298 358 COILED COIL (POTENTIAL).
FT DOMAIN 711 1018 PROTEIN_KINASE.
FT NP_BIND 717 725 ATP (BY SIMILARITY).
FT BINDING 738 738 ATP (BY SIMILARITY).
SQ SEQUENCE 1018 AA; 115027 MW; 1206284ED4E90816 CRC64;
Query Match 48.8%; Score 42; DB 1; Length 1018;
Best Local Similarity 53.8%; Pred. No. 41;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 1 WKKPRNLEEDD 13
Db 646 WKLVEKSNLESD 658
[: : : : :]
[: : : : :]
RESULT 8
ST31_HUMAN
ID ST31_HUMAN STANDARD; PRT; 1019 AA.
AC Q9BXU1; Q9BXH8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serine/threonine protein kinase 31 (EC 2.7.1.37) (Serine/threonine-
DE protein kinase NYD-SPK).
GN STK31.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=21175748; PubMed=11279525;
RA Wang P.J., McCarrey J.R., Yang F., Page D.C.;
RT "An abundance of X-linked genes expressed in spermatogonia.";
RL Nat. Genet. 27:422-426(2001).
RN [2]
RP SEQUENCE OF 24-1019 FROM N.A.
RC TISSUE=Testis;
RA Zhou Z.M.;
RT "Cloning of a new protein kinase gene related to human testis
RT development.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- TISSUE SPECIFICITY: Testis specific.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC -!- SIMILARITY: Contains 1 Tudor domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF285599; AAK31978.1; --
DR EMBL; AF332194; AAK3193.1; --
DR Genew; HGNC:11407; STK31.
DR MIM; 605790; --
DR InterPro; IPR008191; Maternal_tudor.

```

DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase
DE complex (EC 2.3.1.12) (E2) (S complex, 48 kDa subunit).
GN PDHC OR ACEC OR BSO14600.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=90368558; PubMed=1697575;
RA Hemilia H., Palva A., Paulin L., Arvidson S., Palva I.;
RT "Secretory S complex of Bacillus subtilis: sequence analysis and
RT identity to pyruvate dehydrogenase.";
RL J. Bacteriol. 172:5052-5063(1990).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=97124187; PubMed=8969500;
RA Winters P., Caldwell R., Enfield L., Ferrari E.;
RT "The ampS-nprE (124 degrees-127 degrees) region of the Bacillus
RT subtilis 168 chromosome: sequencing of a 27 kb segment and
RT identification of several genes in the area.";
RL Microbiology 142:3033-3037(1996).
RN [3]
RN SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriest L., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Gim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Konigstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Mostl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadale E.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serron P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconis E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tanakashi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambuit R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256(1997).
CC -!- FUNCTION: The pyruvate dehydrogenase complex catalyzes the overall
CC conversion of pyruvate to acetyl-CoA and CO(2). It contains
CC multiple copies of three enzymatic components: pyruvate
CC dehydrogenase (E1), dihydrolipoamide acetyltransferase (E2) and
CC lipoamide dehydrogenase (E3).
CC -!- FUNCTION: THE B.SUBTILIS PDH COMPLEX POSSESSES ALSO BRANCHED-CHAIN

```

```

CC 2-OXOACID DEHYDROGENASE (BCDH) ACTIVITY.
CC -!- CATALYTIC ACTIVITY: Acetyl-CoA + dihydrolipoamide = CoA + S-
CC acetyldihydrolipoamide.
CC -!- COFACTOR: Contains 1 covalently bound lipoyl cofactor.
CC -!- SUBUNIT: Forms a 24-polypeptide structural core with octahedral
CC symmetry.
CC -!- SIMILARITY: Belongs to the 2-oxoacid dehydrogenase family.
CC -!- SIMILARITY: Contains 1 lipoyl-binding domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M57435; AAA62683.1; -.
CC EMBL; AF012285; AAC24934.1; -.
CC EMBL; Z99111; CAB13333.1; -.
CC PIR; D36718; D36718.
CC HSSP; P11961; 2PDE.
CC Subtilist; BG10209; pdhC.
CC InterPro; IPR001078; 2Oxoacid dh.
CC InterPro; IPR000089; Biotin_lipoyl.
CC InterPro; IPR004167; E3_binding.
CC InterPro; IPR003016; Lipoyl_BS.
CC Pfam; PF00198; 2-oxoacid dh; 1.
CC Pfam; PF00364; biotin_lipoyl; 1.
CC Pfam; PF02817; e3_binding; 1.
CC ProDom; PD001115; 2Oxoacid dh; 1.
CC ProSITE; PS00189; LIPOYL; 1.
CC GlycoLysis; Transferase; Acyltransferase; Lipoyl; Complete proteome.
FT INIT MET 0 0 BY SIMILARITY.
FT DOMAIN 1 85 LIPOYL BINDING.
FT DOMAIN 139 181 EI/E3 BINDING.
FT BINDING 42 42 LIPOYL.
FT ACT_SITE 412 412 POTENTIAL.
SQ SEQUENCE 441 AA; 47407 MW; 6B44D47879D134DF CRC64;

Query Match 48.8%; Score 42; DB 1; Length 441;
Best Local Similarity 46.7%; Pred. No. 17;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 WKMKPRNLLEDDYL 15
Db 21 WFVKPNDEVEDDVL 35

RESULT 6
AGUA TRIPE
ID AGUA TRIPE STANDARD; PRT; 847 AA.
AC Q99024;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Alpha-glucuronidase precursor (EC 3.2.1.139) (Alpha-glucosiduronase)
DE (GLU1).
OS Trichoderma reesei (Hypocrea jecorina).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; Hypocrea.
OX NCBI_TaxID=51453;
RN [1]
RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=QM9414 / Rut C-30;
RX MEDLINE=96257277; PubMed=8654984;
RA Margolles-Clark E., Saloheimo M., Siika-Aho M., Penttilae M.;
RT "The alpha-glucuronidase-encoding gene of Trichoderma reesei.";
RL Gene 172:171-172(1996).
CC -!- FUNCTION: RELEASES 4-O-METHYLGLUCURONIC ACID FROM XYLAN.
CC -!- CATALYTIC ACTIVITY: An alpha-D-glucuronoside + H(2)O = an alcohol
CC + D-glucuronate.
CC -!- SUBCELLULAR LOCATION: Secreted.

```

RA Borzym K., Langer I., Beck A., Lenrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu S., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Rhode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrrell B.G., Nurse P.;
 RT "The genome sequence of *Schizosaccharomyces pombe*.";
 RL Nature 415:871-880(2002).
 EN [3]
 RP SEQUENCE OF 224-413 FROM N.A.
 RC STRAIN=968 H90;
 RX MEDLINE=20223868; PubMed=10759889;
 RA Ding D.-Q., Tomita Y., Yamamoto A., Chikashige Y., Haraguchi T.,
 RA Hiraoka Y.;
 RT "Large-scale screening of intracellular protein localization in living
 RT fission yeast cells by the use of a GFP-fusion genomic DNA library.";
 RL Genes Cells 5:169-190(2000).
 CC -!- FUNCTION: INVOLVED IN OXIDATIVE STRESS. TRANSCRIPTION FACTOR THAT
 CC ACTS UPON TRP1 AND CTR1.
 CC
 CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -!- SIMILARITY: In the N-terminal section; belongs to the HSF family.
 CC -!- SIMILARITY: Contains 1 response regulatory domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AB0411768; BAB16722.1; -;
 DR EMBL; Z99168; CAB16301.1; -;
 DR EMBL; AB027943; BAA87247.1; -;
 DR PIR; T39150; T39150.
 DR HSSP; P22121; 2HTS.
 DR GeneDB SPombe; SPAC8C9.14; -;
 DR InterPro; IPR000232; HSF_DNA_bind.
 DR InterPro; IPR002341; HSF_ETs.
 DR InterPro; IPR001789; Response_reg.
 DR Pfam; PF00447; HSF_DNA-bind; 1.
 DR Pfam; PF000072; response_reg; 1.
 DR PRINTS; PR00056; HSFDOMAIN.
 DR ProDom; PD001788; HSF_DNA_bind; 1.
 DR ProDom; PD000039; Response_reg; 1.
 DR SMART; SM00415; HSF; 1.
 DR SMART; SM00448; REC; 1.
 DR PROSITE; PS00434; HSF_DOMAIN; 1.
 DR PROSITE; PS00110; RESPONSE_REGULATORY; 1.
 KW Transcription regulation; Sensory transduction; Nuclear protein;
 KW DNA-binding; Phosphorylation.
 FT DNA_BIND 7 111 BY SIMILARITY.
 FT DOMAIN 369 483 RESPONSE_REGULATORY.
 FT MOD_RES 418 418 PHOSPHORYLATION (BY SIMILARITY).
 FT MISC_FEATURE 418 418
 SQ SEQUENCE 539 AA; 60045 MW; F3B4945C595B2871 CRC64;

 Query Match 50.0%; Score 43; DB 1; Length 539;
 Best Local Similarity 60.0%; Pred. No. 14;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

 QY 1 WKMKPRNLEDDYL 15
 ||:|||||
 DB 364 WKRPRIILVEDDEL 378

 RESULT 5
 ODP2_BACSU
 ID ODP2_BACSU STANDARD; PRT; 441 AA.
 AC P21883;
 DT 01-MAY-1991 (Rel. 18, Created)

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; AF186249; AAF17479.1; -;
DR EMBL; AC005053; AAC79150.1; -; ALT_INIT.
DR EMBL; AC004969; AAD15620.2; -;
DR EMBL; BC011802; AAH11802.1; -;
DR Genbank; HGNC:11378; STEAP.
DR MIM; 604415; -;
DR GO; GO:0005887; C:integral to plasma membrane; TAS.
DR GO; GO:0005911; C:intercellular junction; TAS.
DR GO; GO:0015267; F:channel/pore class transporter activity; TAS.
KW Transmembrane; Antigen.
FT TRANSMEM 71 91 POTENTIAL.
FT TRANSMEM 119 139 POTENTIAL.
FT TRANSMEM 164 184 POTENTIAL.
FT TRANSMEM 218 238 POTENTIAL.
FT TRANSMEM 258 278 POTENTIAL.
FT TRANSMEM 291 311 POTENTIAL.
SQ SEQUENCE 339 AA; 39851 MW; 55443AI70C870387 CRC64;

Query Match 100.0%; Score 86; DB 1; Length 339;
Best Local Similarity 100.0%; Pred. No. 4.8e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WKMKPRNLEDDYL 15
DB 14 WKMKPRNLEDDYL 28
|||||

RESULT 2
RND_ECOLI STANDARD; PRT; 375 AA.
AC P09155;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ribonuclease D (EC 3.1.26.3) (RNase D).
GN RND OR B1804.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
[1]
SEQUENCE FROM N.A., AND SEQUENCE OF 1-6.
RX MEDLINE=88289400; PubMed=3041371;
RA Zhang J., Deutscher M.P.;
RA "Escherichia coli RNase D: sequencing of the rnd structural gene and
RT purification of the overexpressed protein."
RL Nucleic Acids Res. 16:6265-6278(1988).
RN [2]
SEQUENCE FROM N.A.
RP STRAIN=K12 / MG1655;
RC MEDLINE=97426617; PubMed=9278503;
RX Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-VIDES J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RA "The complete genome sequence of Escherichia coli K-12."
RT Science 277:1453-1474(1997).
RL [3]
SEQUENCE FROM N.A.
RP STRAIN=K12;
RX MEDLINE=97251358; PubMed=9097040;
RA Itch T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T.,
RA Isono K., Kasai H., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T.,
RA Saito N., Sampei G., Seki Y., Sivasubraman S., Tagami H.,
RA Takeda J., Takemoto K., Wada C., Yamamoto Y., Horiiuchi T.;
RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome

RT corresponding to the 40.1-50.0 min region on the linkage map.";
RL DNA Res. 3:379-392(1996).
RN [4]
RP SEQUENCE OF 1-38 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=94150456; PubMed=8107670;
RA Fulda M., Heinz E., Wolter F.P.;
RT "The fadD gene of Escherichia coli K12 is located close to rnd at
RT 39.6 min of the chromosomal map and is a new member of the
RT AML-binding protein family."
RL Mol. Gen. Genet. 242:241-249(1994).

CC -!- FUNCTION: CLEAVES MULTIMERIC TENA PRECURSOR AT THE SPACER REGION.
CC -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; X07055; CAA30098.1; -;
DR EMBL; AE000274; AAC74874.1; -;
DR EMBL; D90823; BAA15599.1; -;
DR EMBL; D90824; BAA15608.1; -;
DR EMBL; D90825; BAA15613.1; -;
DR EMBL; X70994; CAA50322.1; -;
DR PIR; S01223; NRECD.

DR EcoGene; EGI0858; rnd.
DR InterPro; IPR002562; 3_5_exonuclease.
DR InterPro; IPR002121; HRDC.
DR InterPro; IPR006292; Ribonuclease D.
DR Pfam; PF01612; 3_5_exonuclease; 1.
DR Pfam; PF00570; HRDC; 1.
DR SMART; SM00474; 3SEXOC; 1.
DR SMART; SM00341; HRDC; 1.
DR TIGRfams; TIGR01388; rnd; 1.
KW Hydrolase; Nuclease; Exonuclease; Complete proteome.
SQ SEQUENCE 375 AA; 42734 MW; AC056D3863712BAE CRC64;

Query Match 52.3%; Score 45; DB 1; Length 375;
Best Local Similarity 63.6%; Pred. No. 4.4;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 WKMKPRNLEE 11
DB 342 WKMKPRNLEE 352
|||:|:|:|

RESULT 3
VANI VIBAN STANDARD; PRT; 193 AA.
ID VANI VIBAN
AC P74945;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Autoinducer synthesis protein vanI.
GN VANI.
OS Vibrio anguillarum (Listonella anguillarum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Listonella.
OX NCBI_TaxID=55601;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=NB10 / Serotype O1;
RC MEDLINE=97284506; PubMed=9139920;
RX Milton D.L., Hardman A., Camara M., Chhabra S.R., Bycroft B.W.,
RA Stewart G.S.A.B., Williams P.;
RA "Quorum sensing in Vibrio anguillarum: characterization of the
RT vanI/vanR locus and identification of the autoinducer N-(3-

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 13:52:02 ; Search time 4.7093 Seconds
(without alignments)
165.853 Million cell updates/sec

Title: US-10-010-667a-2_COPY_14_28
Perfect score: 86
Sequence: 1 WKMKPRNLEEDYL 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	86	100.0	339	1	STEA_HUMAN
2	45	52.3	375	1	RND_ECOLI
3	44	51.2	193	1	VANI_SCHPO
4	43	50.0	539	1	PRR1_VIBAN
5	42	48.8	441	1	ODP2_BACSU
6	42	48.8	847	1	AGUA_TREIR
7	42	48.8	1018	1	ST31_MOUSE
8	42	48.8	1019	1	ST31_HUMAN
9	40.5	47.1	639	1	HTPG_TREPA
10	40.5	47.1	1415	1	ICB4_HSVMG
11	40	46.5	36	1	RET4_CHICK
12	40	46.5	137	1	RET3_XENLA
13	40	46.5	197	1	I18B_HUMAN
14	40	46.5	549	1	SVR_ARCFU
15	40	46.5	553	1	GPT1_CANAL
16	39.5	45.9	607	1	HTPG_FUSNN
17	39	45.3	166	1	VSN1_NOCAR
18	39	45.3	347	1	AKSF_METUA
19	39	45.3	457	1	IRF7_MOUSE
20	39	45.3	639	1	SLY1_DROME
21	39	45.3	656	1	SLY1_DROVI
22	39	45.3	766	1	YU20_DROME
23	39	45.3	1469	1	DP27_CABEL
24	38	44.2	142	1	OM22_HUMAN
25	38	44.2	193	1	LUX1_VIBFI
26	38	44.2	345	1	LEU3_LACIA
27	38	44.2	352	1	YG35_MYCFN
28	38	44.2	370	1	HAO1_HUMAN
29	38	44.2	427	1	ODP2_BACST
30	38	44.2	691	1	VP80_NPVAC
31	38	44.2	725	1	YGL4_YEAST
32	38	44.2	729	1	KEX1_YEAST
33	38	44.2	736	1	YBL7_YEAST

34	38	44.2	839	1	YDDB_HABIN
35	38	44.2	925	1	RNT1_SCHPO
36	38	44.2	968	1	PKD2_HUMAN
37	38	44.2	1371	1	YOL4_CABEL
38	38	44.2	3865	1	HRX_MOUSE
39	38	44.2	3969	1	HRX_HUMAN
40	37.5	43.6	716	1	ENPL_RABIT
41	37.5	43.6	795	1	ENPL_CHICK
42	37.5	43.6	802	1	ENPL_MOUSE
43	37.5	43.6	803	1	ENPL_HUMAN
44	37.5	43.6	804	1	ENPL_BOVIN
45	37	43.0	133	1	HEMI_PYRAP

P45182	haemophilus
Q09820	schizosacch
Q13563	homo sapien
Q02331	caenorhabdi
P55200	mus musculu
Q03164	homo sapien
O18750	oryctolagus
P08110	gallus galli
P08113	mus musculu
P14625	homo sapien
Q95ml8	bos taurus
P37363	pyrrhocoris

ALIGNMENTS

RESULT 1

STEAD_HUMAN STANDARD; PRT; 339 AA.
AC OSUHE8; O95034;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Six transmembrane epithelial antigen of prostate.
GN STEAP OR STEAP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20056277; PubMed=10588738;
RA Hubert R.S., Vivanco I., Chen E., Rastegar S., Leong K.,
RA Mitchell S.C., Madraswala R., Zhou Y., Kuo J., Raitano A.B.,
RA Jakobovits A., Safiran D.C., Afar D.E.H.;
RT "STEAP: a prostate-specific cell-surface antigen highly expressed in
RT human prostate tumors."
RL Proc. Natl. Acad. Sci. U.S.A. 96:14523-14528 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Abu-Threideh J., Stoneking T., Langston Y., Maupin R.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Garcia J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Rodriguez A.C., Griewood J., Schmutz J., Myers R.M.,
RA Schmech A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- TISSUE SPECIFICITY: Highly expressed in prostate tumors.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its

Search completed: July 12, 2004, 14:10:23
Job time : 24.1512 secs

RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of *Shigella flexneri* 2a: insights into pathogenicity
through comparison with genomes of *Escherichia coli* K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441(2002).
DR EMBL: AE015166; AAN43025.1; -.
DR GO: GO:0005737; C:cytoplasm; IEA.
DR GO: GO:0008408; F:3'-5' exonuclease activity; IEA.
DR GO: GO:0003676; F:nucleic acid binding; IEA.
DR GO: GO:0004540; F:ribonuclease activity; IEA.
DR GO: GO:0008033; P:RNA processing; IEA.
DR InterPro: IPR002562; 3.5 exonuclease.
DR InterPro: IPR02121; HRDC.
DR InterPro: IPR006292; Ribonuclease_D.
DR Pfam: PF01612; 3.5 exonuclease; 1.
DR Pfam: PF00570; HRDC; 1.
DR SMART: SM00474; 35EXOC; 1.
DR SMART: SM00341; HRDC; 1.
DR TIGRFAMs: TIGR01388; rnd; 1.
KW Complete proteome.
SQ SEQUENCE 375 AA; 42762 MW; 49FA4111C3552B34 CRC64;

Query Match 52.3%; Score 45; DB 16; Length 375;
Best Local Similarity 53.6%; Pred. No. 18;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 WRMKPRNLEE 11
||:|:|:|
DB 342 WKLKPNLNLE 352

RESULT 13
Q8Q0H4 PRELIMINARY; PRT; 461 AA.
ID Q8Q0H4
AC Q8Q0H4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Conserved protein.
DE MW0162.
OS Methanosarcina mazei (Methanosarcina frisia).
OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2209;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Goel / Go1 / ATCC BAA-199 / DSM 3647 / OCM 88;
RX MEDLINE=22120827; PubMed=12125824;
RA Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
RA Martinez-Arias R., Henne A., Waeber A., Baumer S., Jacobi C.,
RA Brueggemann H., Lienard T., Christmann A., Boemcke M., Steckel S.,
RA Bhattacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA Fritz H.-J., Gottschalk G.;
RT "The genome of *Methanosarcina mazei*: evidence for lateral gene
transfer between Bacteria and Archaea.";
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
DR EMBL: AE013238; AAM29858.1; -.
DR InterPro: IPR002576; Archaeal_ATPase.
DR InterPro: IPR004256; DUF234.
DR Pfam: PF01637; Archaeal_ATPase; 1.
DR Pfam: PF03008; DUF234; 1.
KW Complete proteome.
SQ SEQUENCE 461 AA; 54417 MW; 3D1AE9340AE6A9 CRC64;

Query Match 52.3%; Score 45; DB 17; Length 461;
Best Local Similarity 57.1%; Pred. No. 23;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 WRMKPRNLEEDYL 15
|:|:|:|:|:|
DB 439 KVKGRKNLEDDGFL 452

RESULT 14
Q9LGG0 PRELIMINARY; PRT; 2058 AA.
ID Q9LGG0
AC Q9LGG0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE P0406H10.14 protein.
GN P0406H10.14.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoideae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
clone: P0406H10.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP002524; BAB07964.1; -.
DR Gramene; Q9LGG0; -.
SQ SEQUENCE 2058 AA; 226376 MW; 8A06B6562B8BBD69 CRC64;

Query Match 52.3%; Score 45; DB 10; Length 2058;
Best Local Similarity 58.3%; Pred. No. 1.1e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 MKPRNLEEDDY 14
:|:|:|:|
DB 496 LKKRKNISEDDY 507

RESULT 15
O96654 PRELIMINARY; PRT; 317 AA.
ID O96654
AC O96654;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Activated protein kinase C receptor homolog.
OS Trypanosoma vivax.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ILDat1.2;
RA Djikeng A., Majiwa P.A.;
RT "Trypanosoma vivax homologue of the receptor for activated protein
kinase C.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF100287; AAC72850.1; -.
DR GO: GO:0016301; F:kinase activity; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00400; WD40; 7.
DR PRINTS: PR00320; GPROTEINRPT.
DR ProDom: PD000018; WD40; 4.
DR SMART: SM00320; WD40; 7.
DR PROSITE: PS00678; WD_REPEATS_1; 3.
DR PROSITE: PS00082; WD_REPEATS_2; 5.
DR PROSITE: PS0294; WD_REPEATS_REGION; 1.
KW Kinase; Receptor; Repeat; WD-repeat.
SQ SEQUENCE 317 AA; 34895 MW; B3CFB35D520AE31C CRC64;

Query Match 51.2%; Score 44; DB 5; Length 317;
Best Local Similarity 50.0%; Pred. No. 23;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 WRMKPRNLEEDDY 14
|:|:|:|:|
DB 43 WSDNPNRHAENDY 56

```

Query Match      52.3%; Score 45; DB 16; Length 371;
Best Local Similarity 63.6%; Pred. No. 18;
Matches 7; Conservative 2; Mismatches 0; Gaps 0;

QY      1 WKMKPRNLEE 11
      ||:|:|:|:|
Db      338 WKLKPNLNLE 348

RESULT 10
Q8XDR7 PRELIMINARY; PRT; 375 AA.
AC Q8XDR7;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE RNase D, processes tRNA precursor.
GN RND OR Z2847 OR ECS2513.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=0157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoudis K.,
RA Apodaca J., Anantharanan T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RA "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
RT Nature 409:529-533 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=0157:H7 / RMD 050952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22 (2001).
DR EMBL; AF005403; AAC56793.1; -
DR EMBL; AP002558; BAB35936.1; -
DR PIR; A99943; A99943.
DR PIR; E85791; E85791.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0008408; F:3'-5' exonuclease activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0004540; F:ribonuclease activity; IEA.
DR GO; GO:0008033; P:tRNA processing; IEA.
DR InterPro; IPR002562; 3_5_exonuclease.
DR InterPro; IPR002121; HRDC.
DR InterPro; IPR006292; Ribonuclease_D.
DR Pfam; PF01612; 3_5_exonuclease; 1.
DR Pfam; PF00570; HRDC; 1.
DR SMART; SM00341; HRDC; 1.
DR TIGRfams; TIGR01388; rnd; 1.
KW Complete proteome.
SQ SEQUENCE 375 AA; 42748 MW; 0D0F7D2982602BAF CRC64;

Query Match      52.3%; Score 45; DB 16; Length 375;
Best Local Similarity 63.6%; Pred. No. 18;
Matches 7; Conservative 2; Mismatches 0; Gaps 0;

QY      1 WKMKPRNLEE 11
      ||:|:|:|:|
Db      342 WKLKPNLNLE 352

RESULT 11
Q8FGV1 PRELIMINARY; PRT; 375 AA.
AC Q8FGV1;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Ribonuclease D (EC 3.1.26.3).
GN RND OR C2208.
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=06:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
DR EMBL; AF016761; AAN80667.1; -
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0008408; F:3'-5' exonuclease activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0004540; F:ribonuclease activity; IEA.
DR GO; GO:0008033; P:tRNA processing; IEA.
DR InterPro; IPR002562; 3_5_exonuclease.
DR InterPro; IPR002121; HRDC.
DR InterPro; IPR006292; Ribonuclease_D.
DR Pfam; PF01612; 3_5_exonuclease; 1.
DR Pfam; PF00570; HRDC; 1.
DR SMART; SM00341; HRDC; 1.
DR TIGRfams; TIGR01388; rnd; 1.
KW Hydrolyase; Complete proteome.
SQ SEQUENCE 375 AA; 42718 MW; 10A40D3DC7747AEA CRC64;

Query Match      52.3%; Score 45; DB 16; Length 375;
Best Local Similarity 63.6%; Pred. No. 18;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 WKMKPRNLEE 11
      ||:|:|:|:|
Db      342 WKLKPNLNLE 352

RESULT 12
Q83RI1 PRELIMINARY; PRT; 375 AA.
AC Q83RI1;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE RNase D, processes tRNA precursor.
GN RND OR SFI424.
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Zhang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
```

Query Match 54.7%; Score 47; DB 16; Length 243;
 Best Local Similarity 40.0%; Pred. No. 5.2;
 Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Qy 1 WKMKPRRLNEEDDYL 15
 |||:::|::|::|
 Db 91 WKVRFQKLEDSVI 105

RESULT 6

Q8F081 ID Q8F081 PRELIMINARY; PRT; 475 AA.
 AC Q8F081
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Outer membrane protein OmpA family.
 GN LA3615.
 OS Leptospira interrogans.
 CC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
 OC NCBI_TaxID=173;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
 RA Ren S.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; A5011517; AAN50813.1; -;
 DR InterPro; IPR006664; Bac.OmpA.
 DR Pfam; PF00691; OmpA; 1.
 DR PRINTS; PR01021; OMPADOMAIN.
 DR ProDom; PD000930; OmpA/MotB; 1.
 KW Complete proteome.
 SQ SEQUENCE 475 AA; 55426 MW; 5394208BB7DC717B CRC64;

Query Match 54.7%; Score 47; DB 16; Length 475;
 Best Local Similarity 50.0%; Pred. No. 10;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 WKMKPRRLNEEDDYL 14
 |||:::|::|::|
 Db 68 WKLKFGENVELNEY 81

RESULT 7

Q83878 ID Q83878 PRELIMINARY; PRT; 208 AA.
 AC Q83878
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Conserved domain protein.
 GN EF0345.
 OS Enterococcus faecalis (Streptococcus faecalis).
 CC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
 OC NCBI_TaxID=1351;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=V583 / ATCC 700802;
 RX MEDLINE=22550857; PubMed=12663927;
 RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,
 RA Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
 RA Tettelin H., Dodson R.J., Umayam L., Brinkac L., Bearan M.,
 RA Daugherty S., DeBoy R.T., Durkin S., Kolonay J., Madupu R., Nelson W.,
 RA Vamathevan J., Tran B., Opat J., Hansen T., Shetty J., Khouri H.,
 RA Ustebayrak T., Radune D., Ketchum K.A., Dougherty B.A., Fraser C.M.;
 RT "Role of mobile DNA in the evolution of vancomycin-resistant
 Enterococcus faecalis."
 RL Science 299:2071-2074(2003).
 RL EMBL; AE016947; AAC0208.1; -;
 DR TIGR; EF0345; -;
 KW Complete proteome.
 SQ SEQUENCE 208 AA; 22905 MW; 4050BD05B0CBDA98 CRC64;

Query Match 53.5%; Score 46; DB 16; Length 208;
 Best Local Similarity 61.5%; Pred. No. 6.6;
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 MKPRRLNEEDDYL 15
 |||:::|::|::|
 Db 117 IKPKRYLEEGDYI 129

RESULT 8

Q9UTA5 ID Q9UTA5 PRELIMINARY; PRT; 237 AA.
 AC Q9UTA5
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Conserved hypothetical protein.
 GN SPAC25B8.15C.
 OS Schizosaccharomyces pombe (Fission yeast).
 CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OC NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972h-;
 RA Barrell B.G., Rajandream M.A., McDougall R.C., McLean J., Harris D.;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL133225; CAB61781.1; -;
 DR PIR; T50202; T50202.
 DR GeneDB SPombe; SPAC25B8.15C; -;
 DR InterPro; IPR003827; DUF207.
 DR Pfam; PF02676; DUF207; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 237 AA; 26997 MW; 29A6CF0D0C720997 CRC64;

Query Match 52.9%; Score 45.5; DB 3; Length 237;
 Best Local Similarity 66.7%; Pred. No. 9.3;
 Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

Qy 2 KMKPRRLNEE-DDYL 15
 |||:::|::|::|
 Db 220 RMKPRKLRNMDDYL 234

RESULT 9

Q7UCL0 ID Q7UCL0 PRELIMINARY; PRT; 371 AA.
 AC Q7UCL0
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE RNase D.
 GN RND OR S1539.
 OS Shigella flexneri.
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Shigella.
 OC NCBI_TaxID=623;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
 RX MEDLINE=22590274; PubMed=12704152;
 RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
 RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
 RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
 RA Schwartz D.C., Blattner F.R.;
 RT "Complete genome sequence and comparative genomics of Shigella
 flexneri serotype 2a strain 2457T."
 RL Infect. Immun. 71:2775-2786(2003).
 DR EMBL; AE016983; AAP16920.1; -;
 SQ SEQUENCE 371 AA; 42225 MW; 00577552B66FEA2D CRC64;

```

01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Six-transmembrane epithelial antigen of the prostate.
GN STEAP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; PubMed=11479226;
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA MEDLINE=21371909; PubMed=11479226;
RA Yang D., Holt G.E., Velders M.P., Kwon E.D., Kast W.M.;
RT "Murine six-transmembrane epithelial antigen of the prostate, prostate
RT stem cell antigen, and prostate-specific membrane antigen: prostate-
RT specific cell-surface antigens highly expressed in prostate cancer of
RT transgenic adenocarcinoma mouse prostate mice.";
RL Cancer Res. 61:5857-5860(2001).
DR EMBL; AF297098; AAK83126.1; -.
DR MGD; MGI:1917608; Steap.
DR GO; GO:0016021; C:integral to membrane; IEA.
KW Transmembrane.
SQ SEQUENCE 339 AA; 39318 MW; 4B26A71FF559E84F CRC64;

Query Match 73.3%; Score 63; DB 11; Length 339;
Best Local Similarity 71.4%; Pred. No. 0.011;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 WKMKPRNLEEDDY 14
Db 14 WKMKPKGNLEDDSY 27

RESULT 3
Q92422
ID Q92422 PRELIMINARY; PRT; 339 AA.
AC Q92422;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dudulin.
GN STEAP OR 101000D01RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RA Seru V., Manivet P., Lamblin D., Vaubourdolle M., Kellermann O.,
RA Loric S.;
RT "Prostate and non-prostate expression of dudulin, the mouse ortholog
RT of human STEAP.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY029584; AAK50537.1; -.
DR MGD; MGI:1917608; Steap.
DR GO; GO:0016021; C:integral to membrane; IEA.
SQ SEQUENCE 339 AA; 39109 MW; 32A2C29F2E333BD0 CRC64;

Query Match 73.3%; Score 63; DB 11; Length 339;
Best Local Similarity 71.4%; Pred. No. 0.011;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 WKMKPRNLEEDDY 14
Db 14 WKMKPKGNLEDDSY 27

RESULT 4
Q9CWR7
ID Q9CWR7 PRELIMINARY; PRT; 339 AA.
AC Q9CWR7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE 2410007B19RIK protein.

```

```

GN STEAP OR 2410007B19RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryonic stem cells;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK010437; BAB26938.1; -.
DR MGD; MGI:1917608; Steap.
DR GO; GO:0016021; C:integral to membrane; IEA.
SQ SEQUENCE 339 AA; 39264 MW; 3F7AB9C7520F0968 CRC64;

Query Match 73.3%; Score 63; DB 11; Length 339;
Best Local Similarity 71.4%; Pred. No. 0.011;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 WKMKPRNLEEDDY 14
Db 14 WKMKPKGNLEDDSY 27

RESULT 5
Q87NG5
ID Q87NG5 PRELIMINARY; PRT; 243 AA.
AC Q87NG5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative DNA polymerase III, epsilon subunit.
GN VPI903.
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=RIMD 2210633 / Serotype O3:K6;
RA MEDLINE=22508454; PubMed=12620739;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Neijma M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae.";
RL Lancet 361:743-749(2003).
DR EMBL; AP005079; BAC60166.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0004527; F:exonuclease activity; IEA.
DR InterPro; IPR006055; Exonuclease.
DR Pfam; PF00929; Exonuclease; 1.
KW Complete proteome.
SQ SEQUENCE 243 AA; 27903 MW; 6696A2E2A0291751 CRC64;

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 14:00:58 ; Search time 22.1512 Seconds
(without alignments)
213.658 Million cell updates/sec

Title: US-10-010-667a-2_COPY_14_28

Perfect score: 86

Sequence: 1 WKMKPRNLEDDYL 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL 25:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phase:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*
 - 15: sp_virus:*
 - 16: sp_bacteriap:*
 - 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	67.5	78.5	338	6	Q9GL50 sus scrofa
2	63	73.3	339	11	Q924J9
3	63	73.3	339	11	Q924J2
4	63	73.3	339	11	Q9CWR7
5	47	54.7	243	16	Q87NG5
6	47	54.7	475	16	Q8F081
7	46	53.5	208	16	Q838T8
8	45.5	52.9	237	3	Q9UTAS
9	45	52.3	371	16	Q7UCLO
10	45	52.3	375	16	Q8XDR7
11	45	52.3	375	16	Q8FGV1
12	45	52.3	375	16	Q83R11
13	45	52.3	461	17	Q8Q0H4
14	45	52.3	2058	10	Q9LGG0
15	44	51.2	317	5	Q96654
16	44	51.2	532	10	O80941

076517 trypanosoma
Q9LUQ5 rhizobium m
Q8EVE2 mycoplasma
Q17557 caenorhabdi
Q8M3F6 poeciliopsi
Q8M3F2 poeciliopsi
Q8M3F1 poeciliopsi
Q94HS8 oryza sativ
Q7XGV6 oryza sativ
Q7XSE9 oryza sativ
Q98856 cynops pyrr
Q8B0U2 lettuce big
Q7XKZ3 oryza sativ
Q928F4 listeria in
Q86S8 porcellio d
O61129 crithidia f
Q7Z96 brachydanio
Q9LXJ7 arabidopsis
Q9C6H1 arabidopsis
Q9FJ11 arabidopsis
Q9LXJ6 arabidopsis
Q81972 bacillus ce
Q9AHK8 borrelia bu
Q9VSE1 drosophila
Q8U1Z6 agrobacteri
Q9TYG6 haemonchus
Q9TZ3 haemonchus
Q9GQW8 haemonchus

ALIGNMENTS

RESULT 1

Q9GL50 PRELIMINARY; PRT; 338 AA.
AC Q9GL50;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Six transmembrane endothelial antigen of PAEC.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Nagasaki T., Boulay G., Coupel S., Coulon F., Tesson L.,
RA Heslan J.-M., Soullou J.-P., Charreau B.;
RT "Differential gene expression in endothelial cells during TNF-alpha-
RT and LPS-mediated activation.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF319659; AAG33868.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
KW Transmembrane.
SQ SEQUENCE 338 AA; 39918 MW; ED490E86E067A32B CRC64;

Query Match 78.5%; Score 67.5; DB 6; Length 338;
Matches 13; Conservativity 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 WKMKPRNLEDDYL 15
Db 14 WKMKPRNLEDDYL 27

RESULT 2

Q924J9 PRELIMINARY; PRT; 339 AA.
ID Q924J9
AC Q924J9;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

ABU98390
 ID ABU98390 standard; protein; 258 AA.
 XX
 AC ABU98390;
 XX
 DT 31-JUL-2003 (first entry)
 XX
 DE Novel human gene STEAP-1 variant 8.
 XX
 KW STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
 KW cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
 KW vaccine.
 XX
 OS Homo sapiens.
 XX
 PN WO2003022995-A2.
 XX
 PD 20-MAR-2003.
 XX
 PF 06-SEP-2002; 2002WO-US028371.
 XX
 PR 06-SEP-2001; 2001US-0317840P.
 PR 05-APR-2002; 2002US-0370387P.
 XX
 PA (AGEN-) AGENSYS INC.
 XX
 PI Paris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;
 XX
 DR WPI; 2003-313240/30.
 DR N-PSDB; ACD02604.
 XX
 PT New composition comprising a substance that modulates the status of a
 PT STEAP-1-related protein, useful for treating and detecting cancer.
 XX
 PS Example 2; Fig 2H; 248pp; English.
 XX
 CC The invention describes a composition comprising a substance that
 CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
 CC any of the 15 sequences of 259 amino acids, given in the specification,
 CC or a molecule that is modulated by the protein. where the status of the
 CC cell that expresses the protein is modulated. The compositions, proteins,
 CC polynucleotides and methods are useful for treating and detecting cancer.
 CC The STEAP-1-related proteins are useful for generating cancer vaccines.
 CC The polynucleotides are useful as tools for delineating, with greater
 CC precision, cytogenetic abnormalities in the chromosomal region that
 CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
 CC the amino acid sequence of a variant of human six transmembrane
 CC epithelial antigen of the prostate or STEAP-1
 XX
 SQ Sequence 258 AA;

Query Match 100.0%; Score 192; DB 6; Length 258;
 Best Local Similarity 100.0%; Pred. No. 3.9e-18;
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RRSYRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 34
 DB 185 RRSYRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 218

Search completed: July 12, 2004, 14:07:23
 Job time : 75.7209 secs

CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
CC any of the 15 sequences of 259 amino acids, given in the specification,
CC or a molecule that is modulated by the protein, where the status of the
CC cell that expresses the protein is modulated. The compositions, proteins,
CC polynucleotides and methods are useful for treating and detecting cancer.
CC The STEAP-1-related proteins are useful for generating cancer vaccines.
CC The polynucleotides are useful as tools for delineating, with greater
CC precision, cytogenetic abnormalities in the chromosomal region that
CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
CC the amino acid sequence of a variant of human six transmembrane
CC epithelial antigen of the prostate or STEAP-1
XX
SQ Sequence 258 AA;

Query Match 100.0%; Score 192; DB 6; Length 258;
Best Local Similarity 100.0%; Pred. No. 3.9e-18;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVVRMEI 34
DB 185 RRSRYKLLNWAYQQVQONKEDAWIEHDVVRMEI 218
|||||

RESULT 13
ABU98387

ID ABU98387 standard; protein; 258 AA.

XX AC ABU98387;

XX DT 31-JUL-2003 (first entry)

XX DE Novel human gene STEAP-1 variant 5.

XX KW STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
XX cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
XX KW vaccine.

XX OS Homo sapiens.

XX PN WO2003022995-A2.

XX PD 20-MAR-2003.

XX PF 06-SEP-2002; 2002WO-US028371.

XX PR 06-SEP-2001; 2001US-0317840P.

XX PR 05-APR-2002; 2002US-0370387P.

XX PA (AGEN-) AGENSYS INC.

XX PI Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;

XX WPI; 2003-313240/30.

XX DR N-PSDB; ACDC02601.

XX PT New composition comprising a substance that modulates the status of a
XX STEAP-1-related protein, useful for treating and detecting cancer.

XX Example 2; Fig 2E; 248pp; English.

XX The invention describes a composition comprising a substance that
XX modulates the status of a protein (I) of 340 or 283 amino acids, or of
XX any of the 15 sequences of 259 amino acids, given in the specification,
XX or a molecule that is modulated by the protein, where the status of the
XX cell that expresses the protein is modulated. The compositions, proteins,
XX polynucleotides and methods are useful for treating and detecting cancer.
XX The STEAP-1-related proteins are useful for generating cancer vaccines.
XX The polynucleotides are useful as tools for delineating, with greater
XX precision, cytogenetic abnormalities in the chromosomal region that
XX encodes STEAP-1 that may contribute to the malignant phenotype. This is
XX the amino acid sequence of a variant of human six transmembrane
XX epithelial antigen of the prostate or STEAP-1

SQ Sequence 258 AA;

Query Match 100.0%; Score 192; DB 6; Length 258;
Best Local Similarity 100.0%; Pred. No. 3.9e-18;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVVRMEI 34

DB 185 RRSRYKLLNWAYQQVQONKEDAWIEHDVVRMEI 218
|||||

RESULT 14
ABU98393

ID ABU98393 standard; protein; 258 AA.

XX AC ABU98393;

XX DT 31-JUL-2003 (first entry)

XX DE Novel human gene STEAP-1 variant 11.

XX KW STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
XX cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
XX KW vaccine.

XX OS Homo sapiens.

XX PN WO2003022995-A2.

XX PD 20-MAR-2003.

XX PF 06-SEP-2002; 2002WO-US028371.

XX PR 06-SEP-2001; 2001US-0317840P.

XX PR 05-APR-2002; 2002US-0370387P.

XX PA (AGEN-) AGENSYS INC.

XX PI Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;

XX WPI; 2003-313240/30.

XX DR N-PSDB; ACDC02607.

XX PT New composition comprising a substance that modulates the status of a
XX STEAP-1-related protein, useful for treating and detecting cancer.

XX Example 2; Fig 2K; 248pp; English.

XX The invention describes a composition comprising a substance that
XX modulates the status of a protein (I) of 340 or 283 amino acids, or of
XX any of the 15 sequences of 259 amino acids, given in the specification,
XX or a molecule that is modulated by the protein, where the status of the
XX cell that expresses the protein is modulated. The compositions, proteins,
XX polynucleotides and methods are useful for treating and detecting cancer.
XX The STEAP-1-related proteins are useful for generating cancer vaccines.
XX The polynucleotides are useful as tools for delineating, with greater
XX precision, cytogenetic abnormalities in the chromosomal region that
XX encodes STEAP-1 that may contribute to the malignant phenotype. This is
XX the amino acid sequence of a variant of human six transmembrane
XX epithelial antigen of the prostate or STEAP-1

SQ Sequence 258 AA;

Query Match 100.0%; Score 192; DB 6; Length 258;
Best Local Similarity 100.0%; Pred. No. 3.9e-18;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVVRMEI 34

DB 185 RRSRYKLLNWAYQQVQONKEDAWIEHDVVRMEI 218
|||||

RESULT 15

PA (AGEN-) AGENSYS INC.
 XX Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;
 XX WPI; 2003-313240/30.
 XX
 XX New composition comprising a substance that modulates the status of a
 PT STEAP-1-related protein, useful for treating and detecting cancer.
 XX
 XX Example 53; Page 169; 248pp; English.
 XX
 XX The invention describes a composition comprising a substance that
 CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
 CC any of the 15 sequences of 259 amino acids, given in the specification,
 CC or a molecule that is modulated by the protein, where the status of the
 CC cell that expresses the protein is modulated. The compositions, proteins,
 CC polynucleotides and methods are useful for generating and detecting cancer.
 CC The STEAP-1-related proteins are useful as tools for delineating, with greater
 CC precision, cytogenetic abnormalities in the chromosomal region that
 CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
 CC the amino acid sequence of a variant of human six transmembrane
 CC epithelial antigen of the prostate or STEAP-1
 XX
 XX Sequence 258 AA;
 XX
 XX Query Match 100.0%; Score 192; DB 6; Length 258;
 XX Best Local Similarity 100.0%; Pred. No. 3.9e-18; Mismatches 0; Gaps 0;
 XX Matches 34; Conservative 0; Indels 0; Gaps 0;
 QY 1 RRSRYKLLNWAYQVQVQNKEDAWIEHDVWRMEI 34
 DB 185 RRSRYKLLNWAYQVQVQNKEDAWIEHDVWRMEI 218
 RESULT 12
 ABU98392
 ID ABU98392 standard; protein; 258 AA.
 AC
 AC ABU98392;
 XX
 XX 31-JUL-2003 (first entry)
 XX
 XX Novel human gene STEAP-1 variant 10.
 XX
 XX STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
 KW cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
 KW vaccine.
 XX
 XX Homo sapiens.
 OS
 XX
 XX WO2003022995-A2.
 PN
 XX
 XX 20-MAR-2003.
 PD
 XX
 XX 06-SEP-2002; 2002WO-US028371.
 PF
 XX
 XX 06-SEP-2001; 2001US-0317840P.
 PR
 XX 05-APR-2002; 2002US-0370387P.
 PR
 XX (AGEN-) AGENSYS INC.
 PA
 XX Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;
 XX WPI; 2003-313240/30.
 XX
 XX N-PSDB; ACD02606.
 DR
 XX
 XX New composition comprising a substance that modulates the status of a
 PT STEAP-1-related protein, useful for treating and detecting cancer.
 XX
 XX Example 2; Fig 2J; 248pp; English.
 XX
 XX The invention describes a composition comprising a substance that

XX Homo sapiens.
 OS
 XX
 XX WO2003022995-A2.
 PN
 XX
 XX 20-MAR-2003.
 PD
 XX
 XX 06-SEP-2002; 2002WO-US028371.
 PF
 XX
 XX 06-SEP-2001; 2001US-0317840P.
 PR
 XX 05-APR-2002; 2002US-0370387P.
 PR
 XX (AGEN-) AGENSYS INC.
 PA
 XX Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;
 XX WPI; 2003-313240/30.
 XX
 XX N-PSDB; ACD02603.
 DR
 XX
 XX New composition comprising a substance that modulates the status of a
 PT STEAP-1-related protein, useful for treating and detecting cancer.
 XX
 XX Example 2; Fig 2G; 248pp; English.
 XX
 XX The invention describes a composition comprising a substance that
 CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
 CC any of the 15 sequences of 259 amino acids, given in the specification,
 CC or a molecule that is modulated by the protein, where the status of the
 CC cell that expresses the protein is modulated. The compositions, proteins,
 CC polynucleotides and methods are useful for generating and detecting cancer.
 CC The STEAP-1-related proteins are useful as tools for delineating, with greater
 CC precision, cytogenetic abnormalities in the chromosomal region that
 CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
 CC the amino acid sequence of a variant of human six transmembrane
 CC epithelial antigen of the prostate or STEAP-1
 XX
 XX Sequence 258 AA;
 XX
 XX Query Match 100.0%; Score 192; DB 6; Length 258;
 XX Best Local Similarity 100.0%; Pred. No. 3.9e-18; Mismatches 0; Gaps 0;
 XX Matches 34; Conservative 0; Indels 0; Gaps 0;
 QY 1 RRSRYKLLNWAYQVQVQNKEDAWIEHDVWRMEI 34
 DB 185 RRSRYKLLNWAYQVQVQNKEDAWIEHDVWRMEI 218
 RESULT 11
 ABU98424
 ID ABU98424 standard; protein; 258 AA.
 AC
 AC ABU98424;
 XX
 XX 31-JUL-2003 (first entry)
 XX
 XX STEAP-1 variant 8PLD4 v.2 #1.
 XX
 XX STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
 KW cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
 KW vaccine.
 XX
 XX Homo sapiens.
 OS
 XX
 XX WO2003022995-A2.
 PN
 XX
 XX 20-MAR-2003.
 PD
 XX
 XX 06-SEP-2002; 2002WO-US028371.
 PF
 XX
 XX 06-SEP-2001; 2001US-0317840P.
 PR
 XX 05-APR-2002; 2002US-0370387P.
 PR
 XX

XX SQ Sequence 255 AA;

Query Match 100.0%; Score 192; DB 6; Length 255;
Best Local Similarity 100.0%; Pred. No. 3.9e-18;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSRYKLLNWAYQQVQNKKEDAWIEHDVWRMEI 34
|||
Db 185 RRSRYKLLNWAYQQVQNKKEDAWIEHDVWRMEI 218

RESULT 8

ABU98428

ID ABU98428 standard; protein; 255 AA.

AC ABU98428;

XX XX

DT 31-JUL-2003 (first entry)

XX XX

DE STEAP-1 variant 8P1D4 v.2 #3.

XX XX

KW STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
KW cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
KW vaccine.

XX XX

OS Homo sapiens.

PN WO2003022995-A2.

XX XX

PD 20-MAR-2003.

XX XX

PF 06-SEP-2002; 2002WO-US028371.

XX XX

PR 06-SEP-2001; 2001US-0317840P.
PR 05-APR-2002; 2002US-0370387P.

XX XX

PA (AGEN-) AGENSYS INC.

XX XX

PI Paris M, Ge W, Raitano AB, Challita-Bid PM, Jakobovits A;

XX XX

DR WPI; 2003-313240/30.

DR N-P5DB; AC002609.

XX XX

PT New composition comprising a substance that modulates the status of a
STEAP-1-related protein, useful for treating and detecting cancer.

PS Example 53; Page 172; 248pp; English.

XX XX

CC The invention describes a composition comprising a substance that
modulates the status of a protein (I) of 340 or 283 amino acids, or of
any of the 15 sequences of 259 amino acids, given in the specification,
or a molecule that is modulated by the protein, where the status of the
cell that expresses the protein is modulated. The compositions, proteins,
polynucleotides and methods are useful for generating and detecting cancer.
The STEAP-1-related proteins are useful for treating and detecting cancer.
The polynucleotides are useful as tools for delineating, with greater
precision, cytogenetic abnormalities in the chromosomal region that
encodes STEAP-1 that may contribute to the malignant phenotype. This is
the amino acid sequence of a variant of human six transmembrane
epithelial antigen of the prostate or STEAP-1

XX SQ Sequence 255 AA;

Query Match 100.0%; Score 192; DB 6; Length 255;
Best Local Similarity 100.0%; Pred. No. 3.9e-18;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSRYKLLNWAYQQVQNKKEDAWIEHDVWRMEI 34
|||
Db 185 RRSRYKLLNWAYQQVQNKKEDAWIEHDVWRMEI 218

RESULT 9

ABU98389

ID ABU98389 standard; protein; 258 AA.

AC ABU98389;

XX XX

DT 31-JUL-2003 (first entry)

XX XX

DE Novel human gene STEAP-1 variant 7.

XX XX

KW STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
KW cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
KW vaccine.

PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 6
LENGTH: 197
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)...(28)
US-09-485-632B-6

Query Match 46.5%; Score 40; DB 4; Length 197;
Best Local Similarity 46.7%; Pred. No. 36;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 WKMKPRNLEEDDY 15
DB 167 WYSPRRGLQEQBEL 181

RESULT 7
US-09-247-373B-54
Sequence 54, Application US/09247373B
Patent No. 6168954
GENERAL INFORMATION:
APPLICANT: MCGONIGLE, BRIAN
APPLICANT: O'KEEFE, DANIEL
TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE ENZYMES
FILE REFERENCE: CL-1108-A
CURRENT APPLICATION NUMBER: US/09/247,373B
CURRENT FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 08/924,747
PRIOR FILING DATE: 1997-09-05
NUMBER OF SEQ ID NOS: 56
SOFTWARE: Microsoft Office 97
SEQ ID NO 54
LENGTH: 221
TYPE: PRT
ORGANISM: SOYBEAN
US-09-247-373B-54

Query Match 46.5%; Score 40; DB 3; Length 221;
Best Local Similarity 44.4%; Pred. No. 41;
Matches 8; Conservative 3; Mismatches 3; Indels 4; Gaps 1;

QY 1 WKMK---PRNLEEDDY 14
DB 18 WTLKIDIPYNTBEDRY 35

RESULT 8
US-09-252-991A-31334
Sequence 31334, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196, 136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 31334
LENGTH: 573
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31334

Query Match 46.5%; Score 40; DB 4; Length 573;
Best Local Similarity 70.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 PRNLEEDDY 14
DB 427 PRPNDEDDY 436

RESULT 9
US-09-621-976-5168
Sequence 5168, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Mline Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 5168
LENGTH: 103
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: UNSURE
LOCATION: 68
OTHER INFORMATION: Xaa = Asp,Tyr
US-09-621-976-5168

Query Match 44.2%; Score 38; DB 4; Length 103;
Best Local Similarity 70.0%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 KPRNLEEDD 13
DB 27 KPEELEDDE 36

RESULT 10
US-09-582-379-4
Sequence 4, Application US/09582379
Patent No. 6492176
GENERAL INFORMATION:
APPLICANT: K.U. Leuven Research & Development
TITLE OF INVENTION: Increase of No. 6492176ule number and nitrogen fixation in
TITLE OF INVENTION: Leguminosae
FILE REFERENCE: 001252
CURRENT APPLICATION NUMBER: US/09/582,379
CURRENT FILING DATE: 2000-10-10
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 193
TYPE: PRT
ORGANISM: Vibri0 fischeri
US-09-582-379-4

Query Match 44.2%; Score 38; DB 4; Length 193;
Best Local Similarity 42.9%; Pred. No. 77;
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 WKMKPRNLEEDDY 14
DB 35 WDLVVENNLSDEY 48

RESULT 11
US-08-403-866-2
Sequence 2, Application US/08403866

```
Patent No. 5643779
GENERAL INFORMATION:
APPLICANT: Ehrlich, Stanislaw
APPLICANT: Godon, Jean-Jacques
APPLICANT: Renault, Pierre
TITLE OF INVENTION: Nucleic acid coding for an alpha-acetolactate
TITLE OF INVENTION: synthase from lacticoccus and its applications
NUMBER OF SEQUENCES: 16
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/403,866
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 20747/30
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 263-1600
TELEX: 978450 (WDT)
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 345 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Lactococcus lactis subsp. lactis
INDIVIDUAL ISOLATE: LEUB
US-08-403-866-2

Query Match          44.2%; Score 38; DB 1; Length 345;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY      2 KMKPRNLEDDYL 15
Db      114 KLSPIRVNVDPL 127

RESULT 12
US-09-252-991A-29449
Sequence 29449, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 29449
LENGTH: 464
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
FEATURE:
NAME/KEY: UNSURE
LOCATION: (227), (276)
OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-29449

Query Match          44.2%; Score 38; DB 4; Length 464;
```

```
Best Local Similarity 60.0%; Pred. No. 2e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      5 PRNLEDDY 14
Db      90 PRSLDRDH 99

RESULT 13
US-09-134-000C-4765
Sequence 4765, Application US/09134000C
Patent No. 6617156
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 032796-032
CURRENT APPLICATION NUMBER: US/09/134,000C
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/055,778
PRIOR FILING DATE: 1997-08-15
NUMBER OF SEQ ID NOS: 6812
SOFTWARE: Patentin version 3.1
SEQ ID NO 4765
LENGTH: 540
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-134-000C-4765

Query Match          44.2%; Score 38; DB 4; Length 540;
Best Local Similarity 46.7%; Pred. No. 2.4e+02;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY      1 WKMPPRNLEDDYL 15
Db      23 WFKVPGDTINDDTL 37

RESULT 14
US-09-107-532A-6422
Sequence 6422, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
TITLE OF INVENTION: ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSER: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 893-5007
```

TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 6422:
SEQUENCE CHARACTERISTICS:
LENGTH: 555 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc_feature
LOCATION: (B) LOCATION 1...555
SEQUENCE DESCRIPTION: SEQ ID NO: 6422:
US-09-107-532A-6422

Query Match 44.2%; Score 38; DB 4; Length 555;
Best Local Similarity 46.7%; Pred. No. 2.4e+02;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 WKMKPRRLLEDDYL 15
| : | | : | | | |
Db 30 WPKVKGDTINEDDTL 44

RESULT 15
US-08-651-999A-7
Sequence 7, Application us/08651999A
Patent No. 6031088
GENERAL INFORMATION:
APPLICANT: Stefan Somlo and Toshio Mochizuki
TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE PKD2 GENE
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: AMSTER, ROTHSTEIN & EBENSTEIN
STREET: 90 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH 1.44 MB STORAGE
MEDIUM TYPE: DISKETTE
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/651,999A
FILING DATE: MAY 23, 1996
ATTORNEY/AGENT INFORMATION:
NAME: ELIZABETH A. BOGOSIAN
REGISTRATION NUMBER: 39,911
REFERENCE/DOCKET NUMBER: 96700/395
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-5995
TELEFAX: (212) 286-0854 or 286-0082
TELEX: TWX 710-581-4766
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 968 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: peptide
HYPOTHETICAL: NO
US-08-651-999A-7

Query Match 44.2%; Score 38; DB 3; Length 968;
Best Local Similarity 46.2%; Pred. No. 4.5e+02;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 WKMKPRRLLEDD 13
| : | | : | | | |
Db 293 WKMPNOTADN 305

Search completed: July 12, 2004, 14:12:23
Job time : 10.2442 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 13:51:07 ; Search time 32.9651 Seconds
(without alignments)
128.567 Million cell updates/sec

Title: US-10-010-667a-2_COPY_14_28
Perfect score: 86
Sequence: 1 WKMKPRRLIEDDYLL 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 segs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB	ID	Description
1	86	100.0	15	3	AAV58201	AAV58201 Human STR
2	86	100.0	15	4	AAE02784	AAE02784 Immunogen
3	86	100.0	66	2	AAV12305	AAV12305 Human 5'
4	86	100.0	109	2	AAV12304	AAV12304 Human 5'
5	86	100.0	254	6	ABU98426	ABU98426 STEAP-1 v
6	86	100.0	254	6	ABU98425	ABU98425 STEAP-1 v
7	86	100.0	235	6	ABU98429	ABU98429 STEAP-1 v
8	86	100.0	255	6	ABU98428	ABU98428 STEAP-1 v
9	86	100.0	258	6	ABU98395	ABU98395 Novel hum
10	86	100.0	258	6	ABU98389	ABU98389 Novel hum
11	86	100.0	258	6	ABU98424	ABU98424 STEAP-1 v
12	86	100.0	258	6	ABU98392	ABU98392 Novel hum
13	86	100.0	258	6	ABU98387	ABU98387 Novel hum
14	86	100.0	258	6	ABU98393	ABU98393 Novel hum
15	86	100.0	258	6	ABU98390	ABU98390 Novel hum
16	86	100.0	258	6	ABU98391	ABU98391 Novel hum
17	86	100.0	258	6	ABU98384	ABU98384 Novel hum
18	86	100.0	258	6	ABU98431	ABU98431 STEAP-1 v
19	86	100.0	258	6	ABU98396	ABU98396 Novel hum
20	86	100.0	258	6	ABU98397	ABU98397 Novel hum
21	86	100.0	258	6	ABU98394	ABU98394 Novel hum
22	86	100.0	258	6	ABU98388	ABU98388 Novel hum
23	86	100.0	258	6	ABU98386	ABU98386 Novel hum
24	86	100.0	258	6	ABU98399	ABU98399 Novel hum
25	86	100.0	258	6	ABU98398	ABU98398 Novel hum

ALIGNMENTS

26	86	100.0	282	6	ABU98432	ABU98432 STEAP-1 v
27	86	100.0	282	6	ABU98427	ABU98427 STEAP-1 v
28	86	100.0	282	6	ABU98385	ABU98385 Novel hum
29	86	100.0	339	3	AAV58194	AAV58194 Human STR
30	86	100.0	339	4	AAV01282	AAV01282 P789P anti
31	86	100.0	339	4	AAV69927	AAV69927 Human pro
32	86	100.0	339	4	AAV78845	AAV78845 Human pro
33	86	100.0	339	4	ABU71818	ABU71818 Prostate
34	86	100.0	339	5	ABU95387	ABU95387 Human p78
35	86	100.0	339	5	ABG61813	ABG61813 Prostate
36	86	100.0	339	6	ABU98383	ABU98383 Novel hum
37	86	100.0	339	6	ABU98414	ABU98414 STEAP-1 v
38	86	100.0	339	6	ABU98430	ABU98430 STEAP-1 v
39	86	100.0	339	6	ABR54499	ABR54499 Prostate
40	86	100.0	339	7	ABU63313	ABU63313 Human six
41	86	100.0	339	7	ADB73573	ADB73573 Prostate
42	86	100.0	339	7	ADB14329	ADB14329 Human pro
43	86	100.0	375	4	AAV79829	AAV79829 Human pro
44	86	100.0	375	4	AAE02780	AAE02780 Human six
45	70	81.4	95	2	AAV11840	AAV11840 Human 5'

RESULT 1

AAV58201
ID AAV58201 standard; peptide; 15 AA.

AC AAV58201;
XX

DT 14-MAR-2000 (first entry)
XX

DE Human STRAP-1 peptide 1, corresponding to STRAP-1 residues 14-28.
XX

KM Serpentine transmembrane antigen of the prostate; STRAP-1; prostate;
XX transmembrane domain; type IIa membrane protein; expression; cancer;

KM prostate cancer; bladder cancer; colon cancer; pancreatic cancer;
KW

KW humoral; anticancer vaccine; antibody; detection; diagnosis; prognosis;
KW monitoring; susceptibility; therapeutic inhibitor; drug targeting;

KW recombinant protein.
XX

OS Synthetic.
OS

XX Homo sapiens.
XX

PN WO962941-A2.
XX

PD 09-DEC-1999.
XX

XX 01-JUN-1999; 99WO-US012157.
XX

XX 01-JUN-1998; 98US-0087520P.
XX

XX 30-JUN-1998; 98US-0091183P.
XX

XX (UROG-) UROGENESYS INC.
XX

XX (AFAR) AFAR D E.
XX

XX (HUBER) HUBERT R S.
XX

XX (LEON) LEONG K.
XX

XX (RAIT) RAITANO A B.
XX

XX (SAFF) SAFFRAN D C.
XX

XX Afer DE, Hubert RS, Leong K, Raitano AB, Saffran DC;
XX

XX WPI; 2000-072832/06.
XX

XX Novel proteins useful as diagnostic markers and therapeutic targets,
XX particularly for prostatic cancer.
XX

XX Example 3C; Page 41; 83pp; English.
XX

XX This sequence represents a synthetic peptide (peptide 1) corresponding to
XX residues 14-28 of STRAP-1 (serpentine transmembrane antigen of the
CC

prostate, AAY58194). This peptide was used to raise polyclonal anti-STRAP-1 N-terminus antibodies in sheep. The antibodies were used to probe Western blots of cell lysates from prostate and non-prostate cancer cell lines, and in immunohistochemical analysis of STRAP-1 expression. STRAP-1 is the prototype member of the STRAP Family of proteins (AAY58194-Y58197) which exhibit a high degree of structural conservation, but which show no significant structural homology to known human proteins. The STRAP-1 gene has been localised to chromosome 7p22. STRAP-1 is thought to be a type IIIa membrane protein and is expressed predominantly in prostate cells in normal human tissues. Structurally, STRAP-1 is a 339 amino acid protein characterised by six transmembrane domains and intracellular N- and C-termini, suggesting that it folds in a "serpentine" manner into three extracellular and two intracellular loops. STRAP-1 mRNA and protein expression is maintained at high levels and throughout all stages of prostate cancer. STRAP-1 mRNA and/or protein is also overexpressed in certain other cancers, including bladder, colon, pancreatic and ovarian cancer. The function of the STRAP proteins is not known. They may be ion channels (from the presence of six transmembrane domains, a feature which is shared by certain ion channels) or gap-junction proteins (from immunohistochemical staining). STRAP-1 and STRAP-2 are cell-surface tumour antigens. Immunisation with a STRAP protein induces cellular and humoral immune responses against STRAP-expressing cells. STRAP proteins may be used to identify specific-binding agents, to produce anticancer vaccines and to generate specific antibodies. The antibodies may be used for detection, prognosis, and monitoring of cancers (or susceptibility to cancer), as therapeutic inhibitors or to target therapeutic agents to their site of action. STRAP nucleic acids may be used for recombinant protein production, as diagnostic and prognostic reagents, for identifying STRAP-expressing cells for screening inhibitors of STRAP expression and for therapeutic modulation/inhibition of STRAP expression. Since high levels of STRAP proteins are exposed on the cell surface, they are easily targeted by systemically administered agents, and because they are expressed mainly on prostatic epithelial cells, agents targeted to them should have minimal side effects on other tissues

Sequence 15 AA:

Query Match 100.0%; Score 86; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 8.3e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WKKKPRNLEDDYL 15
Db 1 WKKKPRNLEDDYL 15

RESULT 2
AAE02784
ID AAE02784 standard; peptide; 15 AA.
XX
AC AAE02784;
XX
DT 06-AUG-2001 (first entry)
XX
DE Immunogenic peptide #1 of human STRAP-1.
XX
KW Human; cytostatic; antiproliferative; vaccine; gene therapy;
KW six transmembrane epithelial antigen of the prostate-1; STRAP-1;
KW chromosome 7p22.3; cancer; prostate; colon; bladder; pancreatic; ovarian;
KW lung; serpentine transmembrane antigen; immunogen.
XX
OS Homo sapiens.
XX
PN WO200140276-A2.
XX
PD 07-JUN-2001.
XX
PF 06-DEC-2000; 2000WO-US033040.
XX
PR 06-DEC-1999; 99US-00455486.
XX
PA (UROG-) UROGENESYS INC.
XX

PI Afar DH, Hubert RS, Raitano AB, Saffran DC, Mitchell SC;
PI Paris M, Jakobovits A;
XX
DR WPI; 2001-367804/38.
XX
PT New STRAP (six transmembrane epithelial antigen of the prostate)
PT proteins, expressed in human cancers, useful for detecting and treating
PT cancer.
XX
PS Example 3C; Page 77; 187pp; English.
XX
CC The present invention relates to human six transmembrane epithelial
CC antigen of the prostate (STRAP) protein. STRAP is a member of cell
CC surface serpentine transmembrane antigens. STRAP gene is used in gene
CC therapy, inhibiting the development or progression of a cancer (eg.
CC prostate, colon, bladder, lung, ovarian and pancreatic) expressing STRAP
CC or inhibiting growth or killing cells expressing STRAP in a patient,
CC comprises administering a vaccine composition to the patient, treating a
CC patient with a cancer that expresses STRAP, or inhibiting growth or
CC killing cells expressing STRAP, comprises administering to the patient a
CC vector encoding single chain monoclonal antibody that comprises the
CC variable domains of the heavy and light chains of the monoclonal antibody
CC that specifically binds to STRAP, such that the vector delivers the
CC single chain monoclonal antibody coding sequence to the cancer cells and
CC the encoded single chain monoclonal antibody is expressed
CC intracellularly. The present sequence is an immunogenic peptide of STRAP-
CC 1. STRAP-1 gene is located on chromosome 7p22.3. This peptide is used to
CC immunise sheep for the generation of sheep polyclonal antibodies towards
CC the amino-terminus of anti-STRAP-1
XX

Sequence 15 AA:

Query Match 100.0%; Score 86; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 8.3e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WKKKPRNLEDDYL 15
Db 1 WKKKPRNLEDDYL 15

RESULT 3
AAY12305
ID AAY12305 standard; protein; 66 AA.
XX
AC AAY12305;
XX
DT 17-JUN-1999 (first entry)
XX
DE Human 5' EST secreted protein SEQ ID NO:336.
XX
KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW forensic; gene therapy; chromosome mapping; signal peptide;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; haematopoiesis regulation; tissue growth regulation;
KW reproductive hormone regulation; chemotactic; chemokine; haemostatic;
KW thrombolytic; anti-inflammatory; tumour inhibition.
XX
OS Homo sapiens.
XX
PN WO9906548-A2.
XX
PD 11-FEB-1999.
XX
PF 31-JUL-1998; 98WO-IB001222.
XX
PR 01-AUG-1997; 97US-00905135.
XX
PA (GSEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Iacroleix B;
XX
DR WPI; 1999-153778/13.

DR N-PSDB; AAX41138.
XX
XX New nucleic acids encoding human secreted proteins - obtained from cDNA
PT libraries prepared from e.g. liver, ovary, brain, prostate, kidney, lung,
PT umbilical cord, placenta and colon tissue.
XX
PS Claim 27; Page 677-678; 824pp; English.
XX
CC AAX41094 to AAX41347 represent 5' expressed sequence tags (ESTs) for
CC human secreted proteins, and encode the proteins given in AAY12261 to
CC AAY12514, respectively. The proteins given represent the signal peptide
CC and an N-terminal fragment of a secreted protein. The nucleic acid
CC sequences can be used for producing secreted human gene products. They
CC can also be used to develop products for diagnosis and therapy. The
CC proteins obtained may have cytokine activity, cell
CC proliferation/differentiation activity, haematopoiesis regulating
CC activity, tissue growth regulating activity, reproductive hormone
CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, tumour inhibition activity or other activities. The products
CC can be used in forensic, gene therapy and chromosome mapping procedures.
CC The sequences can also be used for obtaining corresponding promoter
CC sequences. The nucleic acids encoding the signal peptide can be used for
CC directing extracellular secretion of a polypeptide or the insertion of a
CC polypeptide into a membrane, or importing a polypeptide into a cell
XX
SQ Sequence 66 AA;
XX
Query Match 100.0%; Score 86; DB 2; Length 66;
Best Local Similarity 100.0%; Pred. No. 4e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 WKMKPRRNLEDDYL 15
Db 14 WKMKPRRNLEDDYL 28
XX
RESULT 4
AAY12304
ID AAY12304 standard; protein; 109 AA.
XX
AC AAY12304;
XX
DT 17-JUN-1999 (first entry)
XX
XX Human 5' EST secreted protein SEQ ID NO:335.
XX
XX Human, secreted protein; EST; expressed sequence tag; diagnosis;
XX forensic; gene therapy; chromosome mapping; signal peptide;
XX upstream regulatory sequence; cytokine activity; cell proliferation;
XX differentiation; haematopoiesis regulation; tissue growth regulation;
XX reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
XX thrombolytic; anti-inflammatory; tumour inhibition.
XX
XX Homo sapiens.
XX
XX WO906548-A2.
XX
XX PD 11-FEB-1999.
XX
XX PF 31-JUL-1998; 98WO-IB001222.
XX
XX PR 01-AUG-1997; 97US-00905135.
XX
XX (GEST) GENSET.
XX
XX Dumas Milne Edwards J, Duclert A, Lacroix B;
XX
XX WPI; 1999-153778/13.
XX
XX N-PSDB; AAX41137.
XX
XX New nucleic acids encoding human secreted proteins - obtained from cDNA
PT libraries prepared from e.g. liver, ovary, brain, prostate, kidney, lung,

PT umbilical cord, placenta and colon tissue.
XX
XX Claim 27; Page 677; 824pp; English.
XX
XX AAX41094 to AAX41347 represent 5' expressed sequence tags (ESTs) for
CC human secreted proteins, and encode the proteins given in AAY12261 to
CC AAY12514, respectively. The proteins given represent the signal peptide
CC and an N-terminal fragment of a secreted protein. The nucleic acid
CC sequences can be used for producing secreted human gene products. They
CC can also be used to develop products for diagnosis and therapy. The
CC proteins obtained may have cytokine activity, cell
CC proliferation/differentiation activity, haematopoiesis regulating
CC activity, tissue growth regulating activity, reproductive hormone
CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, tumour inhibition activity or other activities. The products
CC can be used in forensic, gene therapy and chromosome mapping procedures.
CC The sequences can also be used for obtaining corresponding promoter
CC sequences. The nucleic acids encoding the signal peptide can be used for
CC directing extracellular secretion of a polypeptide or the insertion of a
CC polypeptide into a membrane, or importing a polypeptide into a cell
XX
SQ Sequence 109 AA;
XX
Query Match 100.0%; Score 86; DB 2; Length 109;
Best Local Similarity 100.0%; Pred. No. 6.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 WKMKPRRNLEDDYL 15
Db 14 WKMKPRRNLEDDYL 28
XX
RESULT 5
ABU98426
ID ABU98426 standard; protein; 254 AA.
XX
AC ABU98426;
XX
XX 31-JUL-2003 (first entry)
XX
XX STEAP-1 variant 8PID4 v.2 #2.
XX
XX STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
XX cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
XX vaccine.
XX
XX Homo sapiens.
XX
XX WO2003022995-A2.
XX
XX PD 20-MAR-2003.
XX
XX PF 06-SEP-2002; 2002WO-US028371.
XX
XX PR 06-SEP-2001; 2001US-0317840P.
XX
XX PR 05-APR-2002; 2002US-0370387P.
XX
XX (AGEN-) AGENSYS INC.
XX
XX Paris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;
XX
XX WPI; 2003-313240/30.
XX
XX New composition comprising a substance that modulates the status of a
PT STEAP-1-related protein, useful for treating and detecting cancer.
XX
XX Example 53; Page 169-170; 248pp; English.
XX
XX The invention describes a composition comprising a substance that
CC modulates the status of a protein (1) of 340 or 283 amino acids, or of
CC any of the 15 sequences of 259 amino acids, given in the specification,
CC or a molecule that is modulated by the protein, where the status of the

CC cell that expresses the protein is modulated. The compositions, proteins,
CC polynucleotides and methods are useful for treating and detecting cancer.
CC The STEAP-1-related proteins are useful for generating cancer vaccines.
CC The polynucleotides are useful as tools for delineating, with greater
CC precision, cytogenetic abnormalities in the chromosomal region that
CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
CC the amino acid sequence of a variant of human six transmembrane
CC epithelial antigen of the prostate or STEAP-1
XX
SQ Sequence 254 AA;

Query Match 100.0%; Score 86; DB 6; Length 254;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WKMKPRRNLEDDYL 15
14 WKMKPRRNLEDDYL 28

RESULT 6

ABU98425
ID ABU98425 standard; protein; 254 AA.

AC ABU98425;

XX 31-JUL-2003 (first entry)

DE STEAP-1 variant 8PID4 v.1 #1.

XX STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
KW cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
XX vaccine.

XX Homo sapiens.

XX W02003022995-A2.

XX 20-MAR-2003.

PF 06-SEP-2002; 2002WO-US028371.

XX 06-SEP-2001; 2001US-0317840P.

PR 05-APR-2002; 2002US-0370387P.

XX (AGEN-) AGENSYS INC.

XX Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;

XX WPI; 2003-313240/30.

PT New composition comprising a substance that modulates the status of a
STEAP-1-related protein, useful for treating and detecting cancer.

XX Example 53; Page 169-170; 248pp; English.

CC The invention describes a composition comprising a substance that
CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
CC any of the 15 sequences of 259 amino acids, given in the specification,
CC or a molecule that is modulated by the protein, where the status of the
CC cell that expresses the protein is modulated. The compositions, proteins,
CC polynucleotides and methods are useful for treating and detecting cancer.
CC The STEAP-1-related proteins are useful for generating cancer vaccines.
CC The polynucleotides are useful as tools for delineating, with greater
CC precision, cytogenetic abnormalities in the chromosomal region that
CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
CC the amino acid sequence of a variant of human six transmembrane
CC epithelial antigen of the prostate or STEAP-1
XX
SQ Sequence 254 AA;

Query Match 100.0%; Score 86; DB 6; Length 254;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WKMKPRRNLEDDYL 15
14 WKMKPRRNLEDDYL 28

RESULT 7

ABU98429
ID ABU98429 standard; protein; 255 AA.

AC ABU98429;

XX 31-JUL-2003 (first entry)

DE STEAP-1 variant 8PID4 v.3 #2.

XX STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
KW cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
XX vaccine.

XX Homo sapiens.

XX W02003022995-A2.

XX 20-MAR-2003.

PF 06-SEP-2002; 2002WO-US028371.

XX 06-SEP-2001; 2001US-0317840P.

PR 05-APR-2002; 2002US-0370387P.

XX (AGEN-) AGENSYS INC.

XX Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;

XX WPI; 2003-313240/30.

PT New composition comprising a substance that modulates the status of a
STEAP-1-related protein, useful for treating and detecting cancer.

XX Example 53; Page 172; 248pp; English.

CC The invention describes a composition comprising a substance that
CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
CC any of the 15 sequences of 259 amino acids, given in the specification,
CC or a molecule that is modulated by the protein, where the status of the
CC cell that expresses the protein is modulated. The compositions, proteins,
CC polynucleotides and methods are useful for treating and detecting cancer.
CC The STEAP-1-related proteins are useful for generating cancer vaccines.
CC The polynucleotides are useful as tools for delineating, with greater
CC precision, cytogenetic abnormalities in the chromosomal region that
CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
CC the amino acid sequence of a variant of human six transmembrane
CC epithelial antigen of the prostate or STEAP-1
XX
SQ Sequence 255 AA;

Query Match 100.0%; Score 86; DB 6; Length 255;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WKMKPRRNLEDDYL 15
14 WKMKPRRNLEDDYL 28

RESULT 8

ABU98428
ID ABU98428 standard; protein; 255 AA.

AC ABU98428;

XX

QY 1 WKMKPRRNLEDDYL 15
 DB 14 WKMKPRRNLEDDYL 28

RESULT 13
 ID ABU98387 standard; protein; 258 AA.
 AC ABU98387;
 XX
 XX
 DT 31-JUL-2003 (first entry)
 DE Novel human gene STEAP-1 variant 5.
 XX
 XX STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
 KW cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
 KW vaccine.
 XX
 OS Homo sapiens.
 PN WO2003022995-A2.
 XX
 XX 20-MAR-2003.
 XX
 PF 06-SEP-2002; 2002WO-US028371.
 XX
 PR 06-SEP-2001; 2001US-0317840P.
 PR 05-APR-2002; 2002US-0370387P.
 XX
 PA (AGEN-) AGENSYS INC.
 XX
 PI Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;
 XX
 DR WPI; 2003-313240/30.
 DR N-PSDB; ACD02601.
 XX
 PT New composition comprising a substance that modulates the status of a
 PT STEAP-1-related protein, useful for treating and detecting cancer.
 XX
 PS Example 2; Fig 2B; 248bp; English.
 XX
 CC The invention describes a composition comprising a substance that
 CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
 CC any of the 15 sequences of 259 amino acids, given in the specification,
 CC or a molecule that is modulated by the protein, where the status of the
 CC cell that expresses the protein is modulated. The compositions, proteins,
 CC polynucleotides and methods are useful for treating and detecting cancer.
 CC The STEAP-1-related proteins are useful for generating cancer vaccines.
 CC The polynucleotides are useful as tools for delineating, with greater
 CC precision, cytogenetic abnormalities in the chromosomal region that
 CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
 CC the amino acid sequence of a variant of human six transmembrane
 CC epithelial antigen of the prostate or STEAP-1
 XX
 SQ Sequence 258 AA;
 Query Match 100.0%; Score 86; DB 6; Length 258;
 Best Local Similarity 100.0%; Pred. No. 1.7e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WKMKPRRNLEDDYL 15
 DB 14 WKMKPRRNLEDDYL 28

RESULT 14
 ID ABU98393 standard; protein; 258 AA.
 AC ABU98393;
 XX
 XX
 DT 31-JUL-2003 (first entry)

XX
 DE Novel human gene STEAP-1 variant 11.
 XX
 KW STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
 KW cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
 KW vaccine.
 XX
 OS Homo sapiens.
 PN WO2003022995-A2.
 XX
 XX 20-MAR-2003.
 XX
 PF 06-SEP-2002; 2002WO-US028371.
 XX
 PR 06-SEP-2001; 2001US-0317840P.
 PR 05-APR-2002; 2002US-0370387P.
 XX
 PA (AGEN-) AGENSYS INC.
 XX
 PI Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;
 XX
 DR WPI; 2003-313240/30.
 DR N-PSDB; ACD02607.
 XX
 PT New composition comprising a substance that modulates the status of a
 PT STEAP-1-related protein, useful for treating and detecting cancer.
 XX
 PS Example 2; Fig 2K; 248bp; English.
 XX
 CC The invention describes a composition comprising a substance that
 CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
 CC any of the 15 sequences of 259 amino acids, given in the specification,
 CC or a molecule that is modulated by the protein, where the status of the
 CC cell that expresses the protein is modulated. The compositions, proteins,
 CC polynucleotides and methods are useful for treating and detecting cancer.
 CC The STEAP-1-related proteins are useful for generating cancer vaccines.
 CC The polynucleotides are useful as tools for delineating, with greater
 CC precision, cytogenetic abnormalities in the chromosomal region that
 CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
 CC the amino acid sequence of a variant of human six transmembrane
 CC epithelial antigen of the prostate or STEAP-1
 XX
 SQ Sequence 258 AA;
 Query Match 100.0%; Score 86; DB 6; Length 258;
 Best Local Similarity 100.0%; Pred. No. 1.7e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WKMKPRRNLEDDYL 15
 DB 14 WKMKPRRNLEDDYL 28

RESULT 15
 ID ABU98390 standard; protein; 258 AA.
 AC ABU98390;
 XX
 XX
 DT 31-JUL-2003 (first entry)
 DE Novel human gene STEAP-1 variant 8.
 XX
 XX STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
 KW cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
 KW vaccine.
 XX
 OS Homo sapiens.
 PN WO2003022995-A2.
 XX
 XX 20-MAR-2003.

XX 06-SEP-2002; 2002WO-US028371.
 PF
 XX
 PR 06-SEP-2001; 2001US-0317840P.
 PR 05-APR-2002; 2002US-0370387P.
 XX
 PA (AGEN-) AGENSYS INC.

PI Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;
 XX
 DR WPI; 2003-313240/30.
 DR N-PSDB; ACD02604.

XX
 PT New composition comprising a substance that modulates the status of a
 PT STAP-1-related protein, useful for treating and detecting cancer.
 XX

PS Example 2; Fig 2H; 248pp; English.
 XX

CC The invention describes a composition comprising a substance that
 CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
 CC any of the 15 sequences of 259 amino acids, given in the specification,
 CC or a molecule that is modulated by the protein, where the status of the
 CC cell that expresses the protein is modulated. The compositions, proteins,
 CC polynucleotides and methods are useful for treating and detecting cancer.
 CC The STAP-1-related proteins are useful for generating cancer vaccines.
 CC The polynucleotides are useful as tools for delineating, with greater
 CC precision, cytogenetic abnormalities in the chromosomal region that
 CC encodes STAP-1 that may contribute to the malignant phenotype. This is
 CC the amino acid sequence of a variant of human six transmembrane
 CC epithelial antigen of the prostate or STAP-1

XX
 SQ Sequence 258 AA;

Query Match 100.0%; Score 86; DB 6; Length 258;
 Best Local Similarity 100.0%; Pred. No. 1.7e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WKMKPRRNLEDDYL 15
 |||||
 Db 14 WKMKPRRNLEDDYL 28

Search completed: July 12, 2004, 14:07:22
 Job time : 33.9651 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 14:00:58 ; Search time 32.4884 Seconds
(without alignments)
213.658 Million cell updates/sec

Title: US-10-010-667A-2_COPY_92_113
Perfect score: 116
Sequence: 1 REVHPLATSHQYFYKIPILV 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP RODENT:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*
15: SP_VIRUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEP:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	108	93.1	338	6 Q9GL50	Q9GL50 sus scrofa
2	94	81.0	339	11 Q924J9	Q924J9 mus musculus
3	94	81.0	339	11 Q924Z2	Q924Z2 mus musculus
4	94	81.0	339	11 Q9CWR7	Q9CWR7 mus musculus
5	68	58.6	489	11 Q8BWB6	Q8BWB6 mus musculus
6	68	58.6	490	4 Q8NFW2	Q8NFW2 homo sapien
7	68	58.6	490	4 Q8IUR7	Q8IUR7 homo sapien
8	50	43.1	172	16 Q8NZR1	Q8NZR1 streptococc
9	43	42.2	764	16 Q7UR83	Q7UR83 rhodospirill
10	43	42.2	172	16 Q93BL7	Q93BL7 streptococc
11	48	41.4	388	16 Q9A0P6	Q9A0P6 streptococc
12	48	41.4	456	4 Q8TDP3	Q8TDP3 homo sapien
13	48	41.4	487	4 Q8TF03	Q8TF03 homo sapien
14	48	41.4	487	4 Q8NFW6	Q8NFW6 homo sapien
15	48	41.4	488	4 Q9NVB5	Q9NVB5 homo sapien
16	48	41.4	488	4 Q86SR6	Q86SR6 homo sapien

17	48	41.4	498	4 Q7Z389	Q7Z389 homo sapien
18	48	41.4	676	13 Q7ZV52	Q7ZV52 brachydanio
19	48	41.4	929	12 Q9QAP4	Q9QAP4 cervid herp
20	47	40.5	172	16 Q8ES61	Q8ES61 oceanobacil
21	47	40.5	233	10 Q7XUP5	Q7XUP5 arabidopsis
22	47	40.5	427	16 Q8R872	Q8R872 thiomanaer
23	47	40.5	429	16 Q8KES8	Q8KES8 chlorobium
24	47	40.5	431	17 Q89V23	Q89V23 bradyrhizob
25	47	40.5	431	17 Q9HMR3	Q9HMR3 halobacteri
26	47	40.5	434	16 Q8FAH9	Q8FAH9 leptospira
27	47	40.5	438	16 Q7TVW8	Q7TVW8 synchococc
28	47	40.5	439	16 Q7UG05	Q7UG05 rhodospirill
29	47	40.5	442	16 Q7VCRO	Q7VCRO prochloroco
30	47	40.5	442	16 Q7TUC3	Q7TUC3 prochloroco
31	47	40.5	442	16 Q7TUC3	Q7TUC3 prochloroco
32	47	40.5	565	3 Q04600	Q04600 saccharomyc
33	47	40.5	568	4 Q96MU3	Q96MU3 homo sapien
34	47	40.5	675	3 Q9UUE5	Q9UUE5 schizosacch
35	47	40.5	688	10 Q7XUG5	Q7XUG5 oryza sativ
36	47	40.5	749	4 Q8NB24	Q8NB24 homo sapien
37	47	40.5	763	10 Q9LMT2	Q9LMT2 oryza sativ
38	47	40.5	763	10 Q9LGD6	Q9LGD6 oryza sativ
39	47	40.5	763	10 Q9FV78	Q9FV78 oryza sativ
40	47	40.5	806	10 Q9LWM2	Q9LWM2 oryza sativ
41	47	40.5	912	12 Q8QV01	Q8QV01 buffalo herp
42	47	40.5	947	12 Q9QAP7	Q9QAP7 bovine herp
43	47	40.5	1552	4 Q8C0F0	Q8C0F0 homo sapien
44	46.5	40.1	437	2 Q93TY7	Q93TY7 shigella fl
45	46.5	40.1	657	16 Q8NZ96	Q8NZ96 streptococc

ALIGNMENTS

```

RESULT 1
ID Q9GL50 PRELIMINARY; PRT; 338 AA.
AC Q9GL50;
DT 01-MAR-2001 (TRENBLREL. 16, Created)
DT 01-MAR-2001 (TRENBLREL. 16, Last sequence update)
DT 01-JUN-2003 (TRENBLREL. 24, Last annotation update)
DE Six transmembrane endothelial antigen of PAEC.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OC NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Nagasaka T., Bouliday G., Coupel S., Coulon F., Tesson L.,
RA Heslan J.-M., Soullion J.-P., Charreau B.;
RT "Differential gene expression in endothelial cells during TNF-alpha-
RT and LPS-mediated activation."
RI Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF19659; AAC33868.1; -.
DR GO: GO:0016021; C:Integral to membrane; IEA.
KW Transmembrane.
KW SQUENCE 338 AA; 39918 MW; ED490E8E067A32B CRC64;
SQ
Query Match 93.1%; Score 108; DB 6; Length 338;
Best Local Similarity 90.9%; Pred. No. 8.1e-10;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 REVHPLATSHQYFYKIPILV 22
DB 91 REVHPLATSHQYFYKIPILV 112
RESULT 2
ID Q924J9 PRELIMINARY; PRT; 339 AA.
AC Q924J9;
DT 01-DEC-2001 (TRENBLREL. 19, Created)
DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)

```

DT 01-JUN-2003 (TREMELREL. 24, Last annotation update)
 DE Six-transmembrane epithelial antigen of the prostate.
 GN STRAP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RX MEDLINE=21371909; PubMed=11479226;
 RA Yang D., Holt G.E., Velders M.P., Kwon E.D., Kast W.M.;
 RT "Murine six-transmembrane epithelial antigen of the prostate, prostate
 stem cell antigen, and prostate-specific membrane antigen: prostate-
 RT specific cell-surface antigen and prostate highly expressed in prostate cancer of
 RT transgenic adenocarcinoma mouse prostate mice.";
 RL Cancer Res. 61:5857-5860(2001).
 DR EMBL; AF297098; AAC83126.1; -.
 DR MGD; MGI:1917608; Steap.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 KW Transmembrane.
 SQ SEQUENCE 339 AA; 39318 MW; 4B26A71FF559E84F CRC64;

Query Match 81.0%; Score 94; DB 11; Length 339;
 Best Local Similarity 77.3%; Pred. No. 1.9e-07;
 Matches 17; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 REVHPLATSHOQFYKIPILV 22
 DB 92 RELIPLVTSREQFYKIPILV 113

RESULT 3
 Q924Z2 PRELIMINARY; PRT; 339 AA.
 AC Q924Z2;
 DT 01-DEC-2001 (TREMELREL. 19, Created)
 DT 01-DEC-2001 (TREMELREL. 19, Last sequence update)
 DT 01-JUN-2003 (TREMELREL. 24, Last annotation update)
 DE Dudlin.
 GN STEAP OR 1010001D01RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Serru V., Nantvet P., Lambdin D., Vaubourgoille M., Kellermann O.,
 RA Loric S.;
 RT "Prostate and non-prostate expression of dudlin, the mouse ortholog
 RT of human STEAP.";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY029584; AAK50537.1; -.
 DR MGD; MGI:1917608; Steap.
 SQ SEQUENCE 339 AA; 39109 MW; 32A2C29F2E333BD0 CRC64;

Query Match 81.0%; Score 94; DB 11; Length 339;
 Best Local Similarity 77.3%; Pred. No. 1.9e-07;
 Matches 17; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 REVHPLATSHOQFYKIPILV 22
 DB 92 RELIPLVTSREQFYKIPILV 113

RESULT 4
 Q9CWR7 PRELIMINARY; PRT; 339 AA.
 AC Q9CWR7;
 DT 01-JUN-2001 (TREMELREL. 17, Created)
 DT 01-JUN-2001 (TREMELREL. 17, Last sequence update)
 DT 01-OCT-2002 (TREMELREL. 22, Last annotation update)
 DE 2410007H19RIK protein.

GN STRAP OR 2410007H19RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Embryonic stem cells;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arai K., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
 RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schiraldi L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carrini P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 DR EMBL; AK010437; BAB26938.1; -.
 DR MGD; MGI:1917608; Steap.
 SQ SEQUENCE 339 AA; 39264 MW; 3F7AB9C7520F0968 CRC64;

Query Match 81.0%; Score 94; DB 11; Length 339;
 Best Local Similarity 77.3%; Pred. No. 1.9e-07;
 Matches 17; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 REVHPLATSHOQFYKIPILV 22
 DB 92 RELIPLVTSREQFYKIPILV 113

RESULT 5
 Q9BWB6 PRELIMINARY; PRT; 489 AA.
 AC Q9BWB6;
 DT 01-MAR-2003 (TREMELREL. 23, Created)
 DT 01-MAR-2003 (TREMELREL. 23, Last sequence update)
 DT 01-JUN-2003 (TREMELREL. 24, Last annotation update)
 DE Weakly similar to tumor suppressor PHFDE.
 GN 4921538B17RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Head;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA The RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 DR EMBL; AK052981; BAC35230.1; -.
 DR MGD; MGI:1921301; 4921538B17RIK.
 DR GO; GO:0006118; P:electon transport; IEA.
 DR InterPro: IPR004455; NADPoxred_P420.
 DR Pfam: PF03807; F420_oxidored; I.
 SQ SEQUENCE 489 AA; 55760 MW; 98CD63D59DDDF24C CRC64;

Query Match 58.6%; Score 68; DB 11; Length 489;
 Best Local Similarity 65.0%; Pred. No. 0.007;

Matches 13; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 REVHPLATSHOQFYKIP1 20
Db 233 RDVHPYARNQSDFKIP1 251

RESULT 6

Q8NFT2 PRELIMINARY; PRT; 490 AA.

AC Q8NFT2;

DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Six-transmembrane epithelial antigen of prostate 2.

GN STEAP2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_Taxid=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Prostate;

RA Porik K.P., Helenius M.A., Viiskorpi T.;

RT "Cloning and characterization of a novel six-transmembrane protein

STEAP2, down-regulated in androgen-independent prostate cancer."

RL Submitted (DEC-2001) to the EMBL/GenBank/DBD databases.

DR EMBL; AA45138; AA04080.1; -

DR Genew; HGNC:17885; STEAP2.

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0006118; P:electron transport; IEA.

DR InterPro; IPR004455; NADPoxred_F420.

DR Pfam; PF03807; F420_oxidored; 1.

KW Transmembrane.

SQ SEQUENCE 490 AA; 55961 MW; F2E9C30CDACCF81 CRC64;

QY 1 REVHPLATSHOQFYKIP1 20
Db 233 RDVHPYARNQSDFKIP1 252

Query Match 58.6%; Score 68; DB 4; Length 490;

Best Local Similarity 65.0%; Pred. No. 0.0071;

Matches 13; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

RESULT 7
Q8IUE7 PRELIMINARY; PRT; 490 AA.

AC Q8IUE7;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE STAMP1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_Taxid=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Prostate;

RX MEDLINE=2229309; PubMed=12095985;

RA Korkmaz K.S., Elbi C.C., Korkmaz C.G., Loda M., Hager G.L.,

SAaetioğlu F.;

RT "Molecular cloning and characterization of STAMP1, a highly prostate

specific six-trans-membrane protein that is overexpressed in prostate

cancer";

RL J. Biol. Chem. 277:36689-36696(2002).

DR EMBL; AY008445; AAC32149.1; -

DR GO; GO:0006118; P:electron transport; IEA.

DR InterPro; IPR004455; NADPoxred_F420.

DR Pfam; PF03807; F420_oxidored; 1.

SQ SEQUENCE 490 AA; 56027 MW; D1971A8480F7E51 CRC64;

Query Match 58.6%; Score 68; DB 4; Length 490;

Best Local Similarity 65.0%; Pred. No. 0.0071;

Matches 13; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 REVHPLATSHOQFYKIP1 20
Db 233 RDVHPYARNQSDFKIP1 252

RESULT 8

Q8NZR1 PRELIMINARY; PRT; 172 AA.

AC Q8NZR1;

DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Hypothetical phage protein spY18_1784.

GN spY18_1784.

OS Streptococcus pyogenes (serotype M18).

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OX NCBI_Taxid=186103;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MGAS8232 / Serotype M18;

RX MEDLINE=21927593; PubMed=11917108;

RA Smoot J.C., Barbhan K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,

RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,

RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,

RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;

RT "Genome sequence and comparative microarray analysis of serotype M18

group A Streptococcus strains associated with acute rheumatic fever

outbreaks";

RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).

DR EMBL; AE010086; AAL98305.1; -

DR GO; GO:0005727; C:extrachromosomal circular DNA; IEA.

DR GO; GO:0003824; F:catalytic activity; IEA.

DR GO; GO:0003677; F:DNA binding; IEA.

DR GO; GO:0008152; P:metabolism; IEA.

DR InterPro; IPR005811; CoA_ligase.

DR InterPro; IPR003115; ParBc.

DR Pfam; PF02195; ParBc; 1.

KW Complete proteome; Hypothetical protein.

SQ SEQUENCE 172 AA; 19672 MW; B581A8BD93B7EA6 CRC64;

Query Match 43.1%; Score 50; DB 16; Length 172;

Best Local Similarity 38.1%; Pred. No. 2.5;

Matches 8; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 2 EVHPLATSHOQFYKIP1V 22
Db 233 EAVEPYAESIKFGEFKIPV 43

RESULT 9

Q7UD83 PRELIMINARY; PRT; 764 AA.

AC Q7UD83;

DT 01-OCT-2003 (TrEMBLrel. 25, Created)

DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Inter-alpha-trypsin inhibitor family heavy chain-related

protein-hypothetical secreted or membrane-associated protein

containing VWFA domain.

GN RB9664.

OS Rhodopirella baltica.

OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;

OC Planctomycetaceae; Pirellula.

OX NCBI_Taxid=117;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=1;

RX MEDLINE=22735913; PubMed=12835416;

RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Hellmann K., Rabus R.,
RA Schlegner H., Anann R., Reinhardt R.,
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RT strain 1.",
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303 (2003).
DR EMBL; BX294150; CAD76394.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 764 AA; 82826 MW; 5A77412DB5AE3311 CRC64;

Query Match 42.2%; Score 49; DB 16; Length 764;
Best Local Similarity 33.3%; Pred. No. 19;
Matches 7; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

Qy 2 EVIHPATSHQOYFYKPIIV 22
Db 183 EYVHPLEHEDRYVFRFPMV 203

RESULT 10
Q938L7 PRELIMINARY; PRT; 172 AA.

AC Q938L7
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical phage associated protein Spym3_1236 (Phage
associated).
GN SPYM3_1236 OR SP50625.
OS Streptococcus pyogenes, and
OC Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314, 198466;
RN
RN SEQUENCE FROM N.A.
RC SPECIES=S.pyogenes; STRAIN=NIH1;
RA Ikebe T., Wada A., Inagaki Y., Sugama K., Tanaka D., Suzuki R.,
RA Katsukawa C., Fujinaga Y., Abe Y., Watanabe H.;
RT "Complete sequence of temperate phage PhiNIH1.1";
RT Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE FROM N.A.
RC SPECIES=S.pyogenes; STRAIN=MGAS315 / Serotype M3;
RX MEDLINE=22133808; PubMed=12122206;
RA Beres S.B., Sylva G.L., Barbhan K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Musser J.M.;
RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
RT phage-encoded toxins, the high-virulence phenotype, and clone
RT emergence";
RT Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083 (2002).

RN
RN SEQUENCE FROM N.A.
RC SPECIES=S.pyogenes; STRAIN=SSI-1 / Serotype M3;
RA Nakagawa I., Kurokawa K., Nakata M., Tomiyasu Y., Yamashita A.,
RA Yamazaki K., Okahashi N., Kawabata S., Yasunaga T., Hattori M.,
RA Hayashi H., Hamada S.;
RT "The genome of invasive Streptococcus pyogenes; a comparative analysis
RT of S. pyogenes SSI-1, SF370 and MGAS8232.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY050245; AAL15065.1; -
DR EMBL; AE014160; AAM79843.1; -
DR EMBL; AB005143; BAC63720.1; -
DR GO; GO:0005727; C:extrachromosomal circular DNA; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR003115; ParBc.
DR Pfam; PF02195; ParBc; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 172 AA; 19484 MW; ADB1E6565EDAE4C CRC64;

Query Match 41.4%; Score 48; DB 16; Length 172;

Best Local Similarity 38.1%; Pred. No. 5.4;
Matches 8; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Qy 2 EVIHPATSHQOYFYKPIIV 22
Db 23 EAVGPVASEIKERGFYPIIV 43

RESULT 11

Q9A0P6 PRELIMINARY; PRT; 388 AA.

AC Q9A0P6
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical phage associated protein.
GN SPY0679.
OS Streptococcus pyogenes.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN
RN SEQUENCE FROM N.A.
RC STRAIN=SP370 / ATCC 700294 / Serotype M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenyon S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes";
RT Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663 (2001).

RL EMBL; AB006521; AKX33638.1; -
DR GO; GO:0005727; C:extrachromosomal circular DNA; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008170; F:N-methyltransferase activity; IEA.
DR GO; GO:0006306; P:DNA methylation; IEA.
DR InterPro; IPR002295; D21N6_mtfase.
DR InterPro; IPR001091; Met_Tfns_CN4.
DR InterPro; IPR002941; N6_M4_Mcase.
DR InterPro; IPR002052; N6_M4_Mcase.
DR InterPro; IPR003115; ParBc.
DR Pfam; PF01555; N6_N4_Mcase; 1.
DR Pfam; PF02195; ParBc; 1.
DR PRINTS; PR00506; D21N6MTRFAS.
DR PRINTS; PR00508; S21N4MTRFAS.
DR SMART; SM00470; ParB; 1.
DR PROSITE; PS00092; N6_MTASE; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 388 AA; 43690 MW; 12587B620606EB7 CRC64;

Query Match 41.4%; Score 48; DB 16; Length 388;
Best Local Similarity 38.1%; Pred. No. 13;
Matches 8; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Qy 2 EVIHPATSHQOYFYKPIIV 22
Db 23 EAVGPVASEIKERGFYPIIV 43

RESULT 12
Q8TDP3 PRELIMINARY; PRT; 456 AA.

AC Q8TDP3
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE PHYRE II.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN
RN SEQUENCE FROM N.A.

RA Wang C., Allay J.A., Steiner M.S.;
RT "Second human member of pHyde family, Human pHyde II."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF262322; AAM08128.1; -
DR GO:GO:0006118; P:electron transport; IEA.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420 oxidored; 1.
DR PROSITE; PS00290; IG_MHC; 1.
SQ SEQUENCE 456 AA; 54430 MW; C5F7C7008D55251E CRC64;

Query Match 41.4%; Score 48; DB 4; Length 456;
Best Local Similarity 36.4%; Pred. No. 16;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 REVHPLATSHQOYFKIPILV 22
Db 233 RDVLPYQESQNKFFKLPVSV 254

RESULT 13
ID Q8TF03 PRELIMINARY; PRT; 487 AA.
AC Q8TF03;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Tumor suppressor pHyde.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Steiner M.S., Allay J.A., Wang C.;
RT "A novel prostate-derived tumor suppressor."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF238864; AAL78206.1; -
DR GO:GO:0006118; P:electron transport; IEA.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420 oxidored; 1.
DR PROSITE; PS00290; IG_MHC; 1.
SQ SEQUENCE 487 AA; 54402 MW; AAFP16053590E6F68 CRC64;

Query Match 41.4%; Score 48; DB 4; Length 487;
Best Local Similarity 36.4%; Pred. No. 17;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 REVHPLATSHQOYFKIPILV 22
Db 233 RDVLPYQESQNKFFKLPVSV 254

RESULT 14
ID Q8NEM6 PRELIMINARY; PRT; 487 AA.
AC Q8NEM6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE pHyde.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lu Y., Beheshti B., Squire J.A., Yang X.J.;
RT "Characterization of a novel apoptosis-inducing gene, hpHyde, that
inhibits prostate cancer cell growth."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY082673; AAM45136.1; -

DR GO:GO:0006118; P:electron transport; IEA.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420 oxidored; 1.
DR PROSITE; PS00290; IG_MHC; 1.
SQ SEQUENCE 487 AA; 54501 MW; E0EP95855C81EDF CRC64;

Query Match 41.4%; Score 48; DB 4; Length 487;
Best Local Similarity 36.4%; Pred. No. 17;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 REVHPLATSHQOYFKIPILV 22
Db 233 RDVLPYQESQNKFFKLPVSV 254

RESULT 15
ID Q9NVB5 PRELIMINARY; PRT; 488 AA.
AC Q9NVB5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein FLJ10829 (Dudulin 2).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isegaki T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
RA Nakamura Y., Nagahara K., Masuno Y., Sasaki N.;
RT "NEBO human cDNA sequencing project."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Serru V., Manivet P., Lenoir C., Eschwege P., Lambdin D.,
RA Vaudourdolle M., Kellermann O., Loric S.;
RT "Dudulin 2, a new tumor antigen expressed in various human tumors."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK001691; BAA1839.1; -
DR EMBL; AY029585; AAK50538.1; -
DR GO:GO:0006118; P:electron transport; IEA.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420 oxidored; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW Hypothetical protein.

SQ SEQUENCE 488 AA; 54616 MW; BC0BCA48335AAD6 CRC64;

Query Match 41.4%; Score 48; DB 4; Length 488;
Best Local Similarity 36.4%; Pred. No. 17;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 REVHPLATSHQOYFKIPILV 22
Db 233 RDVLPYQESQNKFFKLPVSV 254

Search completed: July 12, 2004, 14:10:24
Job time : 33.4884 secs

This Page Blank (uspto)

Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 13:52:02 ; Search time 6.90698 Seconds

(without alignments)
165.853 Million cell updates/sec

Title: US-10-010-667a-2_COPY_92_113
Perfect score: 116
Sequence: 1 REVHPLATSHQYFYKPIIV 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	116	100.0	339	1	STEPA_HUMAN	Q9uhes homo sapien
2	52.5	45.3	465	1	CIXG_HAENI	P4458 h clixg pro
3	49	42.2	177	1	MSRA_HAENI	Q9hg90 halobacteri
4	47.5	40.9	320	1	CYF_FORPU	P51265 porphyra pu
5	46	39.7	485	1	STPA_ARATH	Q9t034 arabidopsi
6	46	39.7	701	1	GLGE_MYCTU	Q10638 mycobacteri
7	45	39.7	1839	1	CYNA_SACKL	P23466 saccharomyc
8	45	38.8	304	1	WR13_ARATH	Q9svb7 arabidopsi
9	45	38.8	2549	1	FRAP_MOUSE	Q29920 mus musculu
10	44.5	38.4	137	1	SPMI_PIG	Q29920 sus scrofa
11	44.5	38.4	308	1	PRB_METAC	Q8eh12 methanosarc
12	44.5	38.4	1176	1	RPE2_DROME	P08266 drosophila
13	44	37.9	811	1	REX1_YEAST	P48743 saccharomyc
14	44	37.9	928	1	VGLB_HSVBC	P17471 bovine hept
15	44	37.9	932	1	VGLB_HSVBC	P17471 bovine hept
16	43.5	37.5	326	1	YUHS_ECOLI	P39370 escherichia
17	43	37.1	337	1	DHAS_VITCH	P23247 vibrio chol
18	43	37.1	373	1	EXT1_ARATH	Q38913 arabidopsi
19	43	37.1	385	1	GBAS_CAEEL	Q30701 caenorhabdi
20	42.5	36.6	299	1	PYRB_ARCEU	Q30130 archaeoglob
21	42	36.2	216	1	MSRA_XYFLA	Q8pdc4 xyfella fas
22	42	36.2	216	1	MSRA_XYFLA	Q8pdc4 xyfella fas
23	42	36.2	316	1	DHAS_VITAM	Q60800 vibrio mimi
24	42	36.2	580	1	PIGO_MOUSE	Q9qyt7 mus musculu
25	42	36.2	1437	1	DPO3_BACCU	P13267 bacillus su
26	42	36.2	2549	1	FRAP_HUMAN	P42345 homo sapien
27	42	36.2	2549	1	FRAP_HUMAN	P42345 homo sapien
28	41.5	35.8	116	1	AON3_PIG	P24020 sus scrofa
29	41.5	35.8	154	1	MTE_ALUMI	P08200 alligator m
30	41	35.3	194	1	COAE_AGRIS	Q8ujc4 agrobacteri
31	41	35.3	196	1	COAE_HELPJ	Q9z112 helicobacte
32	41	35.3	196	1	COAE_HELPJ	Q9z112 helicobacte
33	41	35.3	216	1	MSRA_XANCP	Q8pca6 xanthomonas

34	41	35.3	575	1	METW_YEAST	Q12198 saccharomyc
35	41	35.3	576	1	SBP_CAEEL	Q21950 caenorhabdi
36	41	35.3	623	1	Y014_MYCPN	P75095 mycoplasma
37	41	35.3	635	1	XYNB_BUTRI	P26223 butyrivibri
38	40.5	34.9	309	1	PYRB_METWA	Q8pdx5 methanosarc
39	40.5	34.9	350	1	XYLD_AGRIS	Q8h7yl agrobacteri
40	40.5	34.9	433	1	PEPD_SALDU	Q8s304 salmonella
41	40.5	34.9	492	1	SYNM_YEAST	P25345 saccharomyc
42	40	34.5	97	1	GATC_PROMP	Q7v354 prochloroco
43	40	34.5	115	1	GUAN_RAT	P28902 rattus norv
44	40	34.5	246	1	YNU0_YEAST	P40165 saccharomyc
45	40	34.5	317	1	PIC_LISMO	P34024 listeria mo

ALIGNMENTS

RESULT 1

STEPA_HUMAN STANDARD; PRT; 339 AA.

AC Q9UHE8; 095034; 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Six transmembrane epithelial antigen of prostate.

GN STEAP OR STEAP1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Carnivora; Hominoidea; Homo.

OX NCBI_TaxId=9606;

RN [1]

RP MEDLINE=20056277; PubMed=10598738.

RX MEDLINE=20056277; PubMed=10598738.

RA Hubert R.S., Vivanco I., Chen E., Raetegar S., Jeong K.,

RA Mitchell S.C., Madraswala R., Zhou Y., Kuo J., Rattano A.B.,

RA Jakobovits A., Saffran D.C., Afar D.B.H.;

RT "STEAP: a prostate-specific cell-surface antigen highly expressed in

RT human prostate tumors."

RL Proc. Natl. Acad. Sci. U.S.A. 96:14523-14528 (1999).

RN [2]

RP MEDLINE=22388257; PubMed=12477932.

RX MEDLINE=22388257; PubMed=12477932.

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,

RA Stappleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,

RA Brownstein M.J., Ustin T.B., Toshitsuki S., Carninci R.P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,

RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gumatratte P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Heltan B., Kettelman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,

RA Schermer A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length

RT human and mouse cDNA sequences."

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RI SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC -!- TISSUE SPECIFICITY: Highly expressed in prostate tumors.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AE005045; AAG19555.1; -
DR PIR: G84273; G84273.
DR HSSP: P54149; 1EVA.
DR HAMAP: ME_01401; -; 1.
DR InterPro: IPR002569; PMSR.
DR Pfam: PF01625; PMSR; 1.
DR Prodom: PD003489; PMSR; 1.
DR TIGRFAMs: TIGR00401; msrA; 1.
DR Oxidoreductase; Complete proteome.
FM ACT SITE 12 BY SIMILARITY.
FT SEQUENCE 177 AA; 19202 MW; 337FBD2E1CFDD8EC CRC64;
SQ

Query Match 42.2%; Score 49; DB 1; Length 177;
Best Local Similarity 50.0%; Pred. No. 1.1;
Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 2 EVINPLATSHQGYFYKIP 19
Db 135 ETFYPAEKHKYFKENP 152

RESULT 4
CYF_PORPU STANDARD; PRT; 320 AA.
AC P51265;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Apocytochrome f precursor.
GN PETA.
OS Porphyra purpurea.
OC Chloroplast.
OC Eukaryota; Bangiophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.
OX NCBI_Taxid=2787;
RN NCBI_Taxid=2787;
RP SEQUENCE FROM N.A.
RC STRAIN=Avonport;
RA Reich M.E., Munholland J.;
RT "Complete nucleotide sequence of the Porphyra purpurea chloroplast
genome";
RL Plant Mol. Biol. Rep. 13:333-335(1995).
CC -!- FUNCTION: Translocates protons across the thylakoid membrane and
CC transfers electrons from photosystem II to photosystem I. It
CC receives electrons from the Rieske iron-sulfur protein and passes
CC them to plastocyanin.
CC -!- SUBUNIT: Interacts with plastocyanin and Rieske iron-sulfur
CC protein (By similarity).
CC -!- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane. Membrane-
CC anchored (By similarity).
CC -!- SIMILARITY: Belongs to the cytochrome c family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: U38804; AAC08151.1; -
DR PIR: S73186; S73186.
DR HSSP: P36438; 1HCZ.
DR HAMAP: MF_00610; -; 1.
DR InterPro: IPR002325; Apocyt_F.
DR InterPro: IPR000345; CytC_heme_BS.

DR Pfam: PF01333; Apocytochr F C; 1.
DR PRINTS: PR00610; CYTOCHROME_F.
DR PROSITE: PS00190; CYTOCHROME_C; 1.
DR KEGG: Electon transport; Heme; Chloroplast; Thylakoid; Photosynthesis;
KM Photosystem I; Photosystem II; Transmembrane; Signal.
FT SIGNAL 1 36
FT CHAIN 37 320
FT TRANSMEM 286 305
FT METAL 37 37
FT BINDING 57 57
FT BINDING 60 60
FT METAL 61 61
SQ SEQUENCE 320 AA; 35012 MW; C9B103F0453369C7 CRC64;
SQ

Query Match 40.9%; Score 47.5; DB 1; Length 320;
Best Local Similarity 40.7%; Pred. No. 3.6;
Matches 11; Conservative 4; Mismatches 7; Indels 5; Gaps 1;

QY 1 REVHPL-----ATSHQGYFYKIPLV 22
Db 160 RELIFPILSPDPKXDKQAHFFKYPYIV 186

RESULT 5
SYFA_ARATH STANDARD; PRT; 485 AA.
ID SYFA_ARATH
AC Q9T034; Q8LPR1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Probable phenylalanyl-tRNA synthetase alpha chain (EC 6.1.1.20)
DE (Phenylalanyl-tRNA ligase alpha chain) (PHERS).
GN ATG439280 OR T22F8.180.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_Taxid=3702;
RN NCBI_Taxid=3702;
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083488; PubMed=10617198;
RA Mayer K.F.X., Schneller C., Wandutt R., Murphy G., Voickaert G.,
RA Pohl T., Duesterhoelt A., Stiekema W., Ertian K.-D., Terry N.,
RA Harris B., Ansoorge W., Brandt P., Griwall L.A., Rieger M.,
RA Weichselgartner M., de Simone V., Obermaler B., Maché R., Mueller M.,
RA Kreis M., Deleney M., Puigdomenech P., Watson M., Schmidtnein T.,
RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA Vos P., Hohnsels U., Zimmermann W., Wedler H., Ridley P.,
RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA Van der Schueren J., Grymoprez B., Chuang Y.-Y., Vandenbussche F.,
RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Deftoor E.,
RA Holzner E., Brandt A., Peters S., van Staveren M., Dirke W.,
RA Woolman P., Klein Lankhorst R., Rose M., Hauf J., Koelster T.-H.,
RA Bernisier S., Hempel S., Feldpausch M., Jambert S., Van den Daele H.,
RA De Keyser A., Buyschaert C., Gelen J., Villarroel R., De Clercq R.,
RA Van Montagu M., Rogers J., Cronin A., Quail M.A., Bray-Allen S.,
RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McIay K., Mayes R.,
RA Petter A., Rajandrem M.A., Lyne M., Bems V., Rechmann S.,
RA Borikova D., Blocher H., Scharfe M., Grimm M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fartman B., Grandenath K., Dauner D., Heitzl A.,
RA Neumann S., Argiriou A., Vitale D., Lignout R., Pizarri E.,
RA Messner O., Quigley F., Clabaud G., Muendlein A., Felber R.,
RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA Cheifor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA Gibbons T., Weber N., Vandenbol M., Barques M., Terol J., Torres A.,
RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Jacou D., Uesse T.,
RA Heijnen L., Schwarz S., Scholler P., Heber S., Frances P., Bietke C.,
RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,

RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA Latorreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA Manx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA Kramer U., Pilton L., Mardis E., Dante M., Pepin K., Hillier L.W.,
RA Nelson J., Speleth J., Ryan E., Andrews S., Geisel C., Layman D.,
RA Du H., Alt J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Jordan C.,
RA Ma P., Zhong N., Preston R., Vil D., Shepherd M., Maturo A., Shah R.,
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman U., Tili S.,
RA Granet S., Shohby N., Hasegawa A., Hamed A., Lodi M., Johnson A.,
RA Chen E., Marra M.A., Martensen R., McCombie W.R.,
RT "Sequence and analysis of chromosome 4 of the plant *Arabidopsis*
RT *thaliana*.";
RL Nature 402:769-777(1999).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, Columbia;
RX MEDLINE=22954850; PubMed=14593172;
RA Yamada K., Lim U., Dale J.M., Chen H., Shinn P., Palm C.J.,
RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
RA Arakawa T., Banh U., Banno F., Bowser L., Brooks S.Y., Carrinci P.,
RA Chao Q., Choy N., Bhui A., Goldsmith A.D., Gujal M., Hansen N.P.,
RA Haysahisaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Kannes M.,
RA Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai U.,
RA Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
RA Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,
RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.,
RT "Empirical analysis of transcriptional activity in the *Arabidopsis*
RT genome.";
RL Science 302:842-846(2003).
CC -1- CATALYTIC ACTIVITY: ATP + L-phenylalanine + tRNA (Phe) = AMP +
CC diphosphate + L-phenylalanyl-tRNA (Phe).
CC -1- SUBUNIT: Tetramer of two alpha and two beta chains (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC Phe-tRNA synthetase alpha chain subfamily 2.
CC -1- CAUTION: Ref.1 sequences differ from that shown due to erroneous
CC gene model prediction.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL050351; CAB43643.1; ALT_SEQ.
DR EMBL: AL161594; CAB80591.1; ALT_SEQ.
DR EMBL: AY094460; AAM19830.1; -.
DR EMBL: BT002310; AAN73307.1; -.
DR HSSP: P27001; 1PYS.
DR InterPro: IPR004529; PheS.
DR InterPro: IPR002319; tRNA-synt_2d.
DR InterPro: IPR006195; tRNA_ligase_II.
DR Pfam: PF01409; tRNA-synt_2d; 1.
DR TIGRFAMs: TIGR00468; pheS; 1.
DR PROSITE: PS50862; AA TRNA_LIGASE II; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; ligase; ATP-binding.
SQ SEQUENCE 485 AA; 55792 MW; 9EA0C0360CA399DB CR64;

Query Match 39.7%; Score 46; DB 1; Length 485;
Best Local Similarity 46.7%; Pred. No. 9.9;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 5 HPLATSHOQYFKIP 19
||| ||| :|||:

DB 257 HPAHSDHTFFIKVP 271
RESULT 6
GIGF_MYCTU STANDARD; PRT; 701 AA.
AC 010638;
DT 01-OCT-1996 (Rel. 34, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Putative glucanase gIGF (EC 3.2.1.-)
GN GIGF OR RV1327C OR MT1369 OR MTCV130.13C OR MB1362C.
OS Mycobacterium tuberculosis, and
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773, 1765;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigemeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holtroyd S.,
RA Hornby T., Jagels K., Krogh A., McLean U., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.,
RT "Deciphering the biology of *Mycobacterium tuberculosis* from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Esmolaeva M., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RT "Whole-genome comparison of *Mycobacterium tuberculosis* clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=M.bovis; STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972;
RA Garner T., Eigemeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Dutfoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Hayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrett B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of *Mycobacterium bovis*.";
RT Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
CC -1- FUNCTION: Could be involved in glycogen catabolism (By
CC similarity).
CC -1- SIMILARITY: Belongs to family 13 of glycosyl hydrolases. GIGF
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z73902; GAA98091.1; ALT_INIT.
DR EMBL: AE007010; AAK5633.1; ALT_INIT.
DR EMBL: BX248338; CAD94223.1; -.
DR TIGR: MT1369; -.
DR TubercuList; RV1327C; -.
DR InterPro: IPR006589; Alp_amy1_cat_sub.

DR InterPro: IPR006047; Alpha_amy1_cat.
 DR Pfam: PF00128; alpha-amy1ase; 1.
 DR SMART: SM00642; Aamy; 1.
 DR Hydrolase: Glycosidase; Complete proteome.
 SQ SEQUENCE 701 AA; 78639 MW; EE637BA1DA7D694F CRC64;

Query Match 39.7%; Score 46; DB 1; Length 701;
 Best Local Similarity 46.7%; Pred. No. 15;
 Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 5 HPLATSHQCYFYKIP 19
 DB 353 HPWAHHRQWFTLPL 367

RESULT 7
 ID CYAA_SACKL STANDARD; PRT; 1839 AA.

AC P23466;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Adenylylase cyclase (EC 4.6.1.1) (ATP pyrophosphate-lyase) (Adenylyl cyclase).
 GN CYRL.
 OS Saccharomyces kluyveri (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4934;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91323718; PubMed=1864503;
 RA Young D., O'Neill K., Broek D., Wiegler M.;
 RT "The adenylyl cyclase-encoding gene from Saccharomyces kluyveri";
 RL Gene 102:123-132(1991).
 CC -!- FUNCTION: Plays essential roles in regulation of cellular metabolism by catalyzing the synthesis of a second messenger, cAMP.
 CC -!- CATALYTIC ACTIVITY: ATP = 3',5'-cyclic AMP + diphosphate.
 CC -!- COFACTOR: Binds 1 magnesium ion per subunit (By similarity).
 CC -!- SIMILARITY: Belongs to the adenylyl cyclase class-3 family.
 CC -!- SIMILARITY: Contains 21 leucine-rich (LRR) repeats.
 CC -!- SIMILARITY: Contains 1 PP2C-like domain.
 CC -!- SIMILARITY: Contains 1 Ras-associating domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X56042; CA39513.1; -;
 DR PIR: J01145; OYBYK.
 DR InterPro: IPR001054; G_cyclase.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR003591; LRR_Typ.
 DR InterPro: IPR001932; PP2C-like.
 DR InterPro: IPR000159; RA_domain.
 DR Pfam: PF00211; guanylate_cyc; 1.
 DR Pfam: PF00560; LRR; 15.
 DR Pfam: PF00481; PP2C; 1.
 DR SMART: SM00044; CYCC; 1.
 DR SMART: SM00352; LRR_Typ; 2.
 DR SMART: SM00352; PP2C; 1.
 DR SMART: SM00314; RA; 1.
 DR PROSITE: PS50125; GUANYLATE_CYCLASES_2; 1.
 DR PROSITE: PS50200; RA; 1.
 KW Lysase; Repeat; Leucine-rich repeat; cAMP biosynthesis; Metal-binding; Magnesium.
 FT DOMAIN 494 574 RAS-ASSOCIATING.
 FT REPEAT 632 655 LRR 1.

FT REPEAT 659 679 LRR 2.
 FT REPEAT 680 702 LRR 3.
 FT REPEAT 703 724 LRR 4.
 FT REPEAT 726 748 LRR 5.
 FT REPEAT 749 771 LRR 6.
 FT REPEAT 773 793 LRR 7.
 FT REPEAT 794 818 LRR 8.
 FT REPEAT 820 833 LRR 9.
 FT REPEAT 834 857 LRR 10.
 FT REPEAT 859 879 LRR 11.
 FT REPEAT 880 903 LRR 12.
 FT REPEAT 905 926 LRR 13.
 FT REPEAT 927 949 LRR 14.
 FT REPEAT 951 974 LRR 15.
 FT REPEAT 1004 1025 LRR 16.
 FT REPEAT 1026 1048 LRR 17.
 FT REPEAT 1050 1072 LRR 18.
 FT REPEAT 1074 1096 LRR 19.
 FT REPEAT 1101 1124 LRR 20.
 FT REPEAT 1135 1150 LRR 21.
 FT DOMAIN 1185 1440 PP2C-LIKE.
 FT DOMAIN 1441 1839 CATALYTIC (BY SIMILARITY).
 FT METAL 1488 1488 MAGNESIUM (BY SIMILARITY).
 FT METAL 1531 1531 MAGNESIUM (BY SIMILARITY).
 SQ SEQUENCE 1839 AA; 206895 MW; 86A69BCH1F2733CB CRC64;

Query Match 39.7%; Score 46; DB 1; Length 1839;
 Best Local Similarity 66.7%; Pred. No. 42;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 VHHPLATSHQCY 14
 DB 589 VHHPLATSHQCY 600

RESULT 8
 ID WR13_ARATH STANDARD; PRT; 304 AA.
 AC Q9SVB7;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Probable WRKY transcription factor 13 (WRKY DNA-binding protein 13).
 GN WRKY13 OR AR4G39410 OR P23K16.40.
 GN Arabidopsis thaliana (Mouse-ear cress).
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 CC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia; TISSUE=flower;
 RA Ulker B., Kushnir S., Somsich I.E.;
 RT "Arabidopsis thaliana transcription factor WRKY13";
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=20083488; PubMed=10617198;
 RA Mayer K.F.X., Schueller C., Mambutt R., Murphy G., Volckaert G.,
 RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terry N.,
 RA Harris B., Ansoorge W., Brandt P., Grivell L.A., Rieger M.,
 RA Weichselgartner M., de Simone V., Obermaier B., Mache R., Weiler M.,
 RA Kreis M., Delseny M., Puldomenech P., Watson M., Schmidheini T.,
 RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
 RA Vos P., Hohnsels J., Zimmermann W., Wedler H., Ridley P.,
 RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
 RA Van der Schuren J., Gyomai B., Chang Y.-J., Vandenbussche F.,
 RA Breken M., Weltjens I., Voel M., Bastiaens I., Aert R., Defoer E.,
 RA Weltzienegger T., Bothe G., Kampsperger U., Hilbert H., Braun M.,
 RA Holzner E., Brandt A., Peters S., van Staveren W., Dirkse W.,
 RA Moeljan P., Klein Iankhorst R., Rose M., Hauf J., Koetter P.,
 RA Berner S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,

```

DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE FRAP-rapamycin associated protein (FRAP).
GN FRAP OR FRAP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALE/c;
RA Bliskovsky V., Mock B.;
RT "Positional cloning of mouse plasmacytoma susceptibility gene.";
RL Submitted (MAY-1999) to the EMBL/Genbank/DBCP databases
CC -! FUNCTION: ACTS AS THE TARGET FOR THE CELL-CYCLE ARREST AND
CC IMMUNOSUPPRESSIVE EFFECTS OF THE FKBP12-RAPAMYCIN COMPLEX.
CC -! SIMILARITY: Belongs to the PI3/P14-kinase family.
CC -! SIMILARITY: Contains 8 HEAT repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/isb-sib.ch).
CC or send an email to license@isb-sib.ch.
CC -----
CC EMBL; AF152838; AAF73196.1; -.
DR HSSP; P42345; IFAF.
DR MGP; MG1:1928394; Frap1.
DR GO: GO:0007281; P:germ-cell development; IDA.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR003151; FAT.
DR InterPro; IPR003152; FATC.
DR InterPro; IPR000357; HEAT.
DR InterPro; IPR000403; PI3_P14_kinase.
DR InterPro; IPR008940; Pretyl1_trans.
DR Pfam; PF02259; FAT; 1.
DR Pfam; PF02260; FATC; 1.
DR Pfam; PF00454; PI3_P14_kinase; 1.
DR SMART; SM00146; PI3KC; 1.
DR PROSITE; PS00915; PI3_4_KINASE_1; 1.
DR PROSITE; PS00916; PI3_4_KINASE_2; 1.
DR PROSITE; PS50290; PI3_4_KINASE_3; 1.
DR PROSITE; PS50077; HEAT_REPEAT; FALSE_NEG.
DR Transferase; Kinase; Repeat.
FT REPEAT 16
FT REPEAT 53
FT REPEAT 650
FT REPEAT 859
FT REPEAT 897
FT REPEAT 988
FT REPEAT 1069
FT REPEAT 1109
FT REPEAT 1150
FT REPEAT 1382
FT REPEAT 1933
FT REPEAT 2182
FT DOMAIN 2549
SQ SEQUENCE 2549 AA; 288734 MM; CQ26B36FE861BC6 CXC64;

Query Match 38.8%; Score 45; DB 1; Length 2549;
Best Local Similarity 50.0%; Pred. No. 87;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 REVHPLATSHOOQFY 16
Db 1683 RQLDHPLEPTAHPOVY 1698

RESULT 10
SPMI_PIG STANDARD; PRT; 137 AA.
AC 028920;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)

```

```

RA SFRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA FitzHugh W., Calvo S., Engels R., Smirnov S., Alnoor D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann N., Dearellano K., Johnson R.,
RA Linton L., McEwan P., McKernan K., Talamas J., Titrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guse A.M.,
RA Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,
RA Leish J.A., Li W., Liu Y., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT and physiological diversity.";
RL Genome Res. 12:532-542(2002).
CC -!- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-aspartate = phosphate
CC + N-carbamoyl-L-aspartate.
CC -!- PATHWAY: Pyrimidine biosynthesis; second step.
CC -!- SUBUNIT: Heterooligomer of catalytic and regulatory chains.
CC -!- SIMILARITY: Belongs to the ATCase/OTCase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/isb-sib.ch).
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE011170; JAM07842.1; -.
DR HAMAP; MF_00001; -; 1.
DR InterPro; IPR006130; Asp/Orn Cotransf.
DR InterPro; IPR002082; Asp carbmltransf.
DR InterPro; IPR006131; OTCase_O.
DR InterPro; IPR006132; OTCase_P.
DR Pfam; PF00185; OTCase; 1.
DR Pfam; PF02729; OTCase_N; 1.
DR PRINTS; PR00100; AOTCASE.
DR TIGRFAMs; TIGR00670; asp carb tr; 1.
DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.
DR Pyrimidine biosynthesis; Transferase; complete proteome.
KW KW SEQUENCE 308 AA; 34644 MW; 33282B438E28595C CRC64;
QY 2 EVIHPD-----ATSH---QQFYKIPK 20
Db 262 KVLHPLEPRVNEIAPEVDATPHACFEQAFGCVPI 295

RESULT 12
ID_RPB2_DROME STANDARD; PRT; 1176 AA.
AC_P08266; Q04155; Q95027; Q9VFW1;
DT 01-AUG-1988 (Rel. 08, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DB DNA-directed RNA polymerase II (EC 2.7.7.6)
DB (RNA polymerase II subunit 2).
GN RPL140 OR CG3180.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RC SEQUENCE FROM N.A.
TC TISSUE=Embryo;
XA MEDLINE=88011299; PubMed=3116266;
FA Falkenburg D., Dworniczak B., Faust D.M., Bautz E.K.F.;
```

"RNA polymerase II of Drosophila. Relation of its 140,000 Mr subunit to the beta subunit of Escherichia coli RNA polymerase.";
 RT J. Mol. Biol. 195:929-937(1987).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortan J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,
 RA Abiri J.F., Agayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
 RA Bailey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahike C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durkin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalish F., Karpen G.H., Ke Z., Kemison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleby J.E.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E.C., Spieding A.C., Stapleton M., Strong M., Sun E.,
 RA Svitskas R., Tector C., Turner K., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley; TISSUE=Embryo;
 RX MEDLINE=242606; PubMed=12537569;
 RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
 RA George R.A., Guarin H., Kronmiller B., Pacleby J.M., Park S., Wan K.H.,
 RA Rubin G.M., Celinker S.E.;
 RT "A Drosophila full-length cDNA resource.";
 RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
 [4]
 RP SEQUENCE OF 1-69 FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=91276237; PubMed=1905256;
 RA Sitrler S., Oldenburg I., Peterson G., Bantz E.K.F.;
 RT "Analysis of the promoter region of the housekeeping gene DmP140 by
 RT sequence comparison of Drosophila melanogaster and Drosophila
 RT virilis.";
 RL Gene 100:155-162(1991).
 CC -I- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription
 CC of DNA into RNA using the four ribonucleoside triphosphates as
 CC substrates.
 CC -I- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC {RNA} (N).
 CC -I- SUBUNIT: EACH CLASS OF RNA POLYMERASE IS ASSEMBLED FROM 9 TO 14
 CC DIFFERENT POLYPEPTIDES. RNA POLYMERASE II CONSISTS OF 10 DIFFERENT
 CC SUBUNITS. THIS SUBUNIT IS THE SECOND LARGEST COMPONENT OF RNA
 CC POLYMERASE II.
 CC -I- SUBCELLULAR LOCATION: Nuclear.

CC -I- MISCELLANEOUS: Three distinct zinc-containing RNA polymerases are
 CC found in eukaryotic nuclei: polymerase I for the ribosomal RNA
 CC precursor, polymerase II for the mRNA precursor, and polymerase
 CC III for 5S and tRNA genes.
 CC -I- SIMILARITY: Belongs to the RNA polymerase beta chain family.
 CC -I- CAUTION: Ref.1 sequence differs from that shown due to a
 CC frameshift in position 43 leading to an erroneous gene model
 CC prediction.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, X05709; CAA29180.2; ALT FRAME.
 DR EMBL, AE003703; AAF55024.1; -.
 DR EMBL, EF003265; AAO25022.1; -.
 DR EMBL, M62972; AAA28476.1; -.
 DR PIR, A27826; A27826.
 DR Flybase, FBgn003276; Rp11140.
 DR GO, GO:0005665; C:DNA-directed RNA polymerase II, core complex; NAS.
 DR GO, GO:000899; P:DNA-directed RNA polymerase activity; NAS.
 DR GO, GO:0006366; P:transcription from Pol II promoter; NAS.
 DR InterPro, IPR007121; RNA pol B.
 DR InterPro, IPR007644; RNA pol_Rpb2_1.
 DR InterPro, IPR007642; RNA pol_Rpb2_2.
 DR InterPro, IPR007645; RNA pol_Rpb2_3.
 DR InterPro, IPR007646; RNA pol_Rpb2_4.
 DR InterPro, IPR007647; RNA pol_Rpb2_5.
 DR InterPro, IPR007120; RNA pol_Rpb2_6.
 DR InterPro, IPR007641; RNA pol_Rpb2_7.
 DR Pfam, PF04563; RNA pol_Rpb2_1; 1.
 DR Pfam, PF04561; RNA pol_Rpb2_2; 1.
 DR Pfam, PF04565; RNA pol_Rpb2_3; 1.
 DR Pfam, PF04566; RNA pol_Rpb2_4; 1.
 DR Pfam, PF04567; RNA pol_Rpb2_5; 1.
 DR Pfam, PF00562; RNA pol_Rpb2_6; 1.
 DR Pfam, PF04560; RNA pol_Rpb2_7; 1.
 DR PROSITE, PS01166; RNA_POL_BETA; 1.
 DR TRANSFASER, DNA-directed RNA polymerase; Transcription; zinc;
 DR TRANSFASER, Zinc-finger; Metal-binding; Nuclear protein.
 DR ZINC FINGER, 1121 1142 C4-TYPE (POTENTIAL).
 FT ZN FING 1121 1142 A->R (IN REF.1).
 FT CONFILCT 72 72
 FT CONFILCT 666 667 ID->MY (IN REF.1).
 SQ SEQUENCE 1176 AA; 134042 MW; 224821B35BED7F0 CRC64;
 Query Match 38.4%; Score 44.5; DB 1; Length 1176;
 Best local Similarity 45.5%; Pred. No. 45;
 Matches 10; Conservative 5; Mismatches 6; Indels 1; Gaps 1;
 QY 2 EVIHPPLATSHQGY-KIPILV 22
 DB 136 EGIIDVETDHCXKTFKIPIML 157
 ID REF1_YEAST STANDARD; PRT; 811 AA.
 AC P48743;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE RFX1-like DNA-binding protein RFX1.
 GN RFX1 OR CRT1 OR YLR176C OR L9470.18.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.

CC STRAIN=S288c / AB972;
RX MEDLINE=97313267; PubMed=9169871;
RA Johnston M., Hillier L., Riles L., Albertmann K., Andre B., Anserge W.,
RA Benes V., Brueckner M., Delius H., Dubois E., Duesterhoef A.,
RA Etian K.-D., Floeth M., Goffeau A., Hebling U., Heumann K.,
RA Heuss-Netzel D., Hilbert H., Hilger F., Kline K., Koetter P.,
RA Louis E.-J., Messerdy F., Mewes H.-W., Miosga T., Moestl D.,
RA Mueller-Auer S., Newtich U., Obermaier B., Piravandi E., Pohl T.M.,
RA Portetelle D., Purnelle B., Reichmann S., Rieger M., Rinke M., Rose M.,
RA Scharte M., Scherens M., Scholler P., Schweger C., Schwarz S.,
RA Underwood A.P., Urestarazu L.A., Vanderbol M., Verhasselt P.,
RA Vierendeels F., Voet M., Volckaert G., Voss H., Wambolt R., Wedler E.,
RA Weller H., Zimmermann F.K., Zollner A., Hant J., Hobetsel J.D.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
RL Nature 387:87-90(1997).
CC -1- SIMILARITY: Belongs to the RFX family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and that this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC
CC EMBL; U17246; AAB67470.1; -.
DR GermOnline; 14238; -.
DR SGD; S0004166; RFX1.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0016566; F:specific transcriptional repressor activity; IDA.
DR GO; GO:0000122; P:negative regulation of transcription from P. . .; IDA.
DR InterPro; IPR003150; RFX_DNA_binding.
DR Pfam; PF02257; RFX_DNA_binding; 1.
KW DNA-binding.
SQ SEQUENCE 811 AA; 90583 MW; 116A88B7DD64FBF0 CRC64;

Query Match 37.9%; Score 44; DB 1; Length 811;
Best local Similarity 60.0%; Pred. No. 36;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Cy 5 HPLATSHOOQFYKTP 19
|||:::|||||
Db 496 HPLITSYKLDPFKIP 510

RESULT 14
VGLB_HSVBP STANDARD; PRT; 928 AA.
ID VGLB_HSVBP
AC P17471;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein I precursor (Glycoprotein GTP-6) (Glycoprotein 11A)
DE (Glycoprotein 16) (Glycoprotein G130).
GN GI.
OS Bovine herpesvirus type 1 (strain P8-2).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicelloviruses.
OX NCBI_Taxid=10324;
RN 1)
RP SEQUENCE FROM N.A.
RX MEDLINE=89020821; PubMed=2845660;
RA Misra V., Nelson R., Smith M.;
RT "Sequence of a bovine herpesvirus type-1 glycoprotein gene that is
RT homologous to the herpes simplex gene for the glycoprotein gB";
RL Virology 166:542-549(1988).
CC -1- SUBUNIT: Dimer, probably linked by disulfide bonds.
CC -1- SIMILARITY: Belongs to the herpesviruses glycoprotein B family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by non-commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M23257; AAA46013.1; -.
DR PIR; A31166; VGBEBG.
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 67
FT CHAIN 68 928
FT DOMAIN 68 780
FT TRANSMEM 781 801
FT TRANSMEM 804 824
FT DOMAIN 825 928
FT CARBOHYD 105 105
FT CARBOHYD 153 153
FT CARBOHYD 442 442
FT CARBOHYD 484 484
FT CARBOHYD 579 579
FT CARBOHYD 637 637
FT CARBOHYD 703 703
FT CONFLICT 409 409
FT CONFLICT 673 673
SQ SEQUENCE 928 AA; 102177 MW; B47982224FCD769D CRC64;
Query Match 37.9%; Score 44; DB 1; Length 928;
Best Local Similarity 40.0%; Pred. No. 42;
Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
Gy 1 REVHPLATSHQYF 15
|:::|:::|
VGLB_HSVBC STANDARD; PRT; 932 AA.
ID VGLB_HSVBC
AC P12640;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein 1 precursor (Glycoprotein GVP-6) (Glycoprotein 11A)
DE (Glycoprotein 16) (Glycoprotein G130) (Glycoprotein B).
GN GI OR UL27.
OS Bovine herpesvirus type 1 (strain Cooper).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10323;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88300884; PubMed=2841484;
RA Whitebeck J.C., Bello L.J., Lawrence W.C.;
RT "Comparison of the bovine herpesvirus 1 gI gene and the herpes
RL J. Virol. 62:3319-3327(1988)."
CC -!- SUBUNIT: Dimer, probably linked by disulfide bonds.
CC -!- SIMILARITY: Belongs to the herpesviruses glycoprotein B family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by non-commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M21474; AAA46055.1; -.
CC EMBL; Z78205; CAB01598.1; -.
CC EMBL; AJ004801; CA06106.1; -.
CC PIR; A28877; VGBEBG.

```

DR InterPro: IPR00234; Glycoprot_B.
 DR Pfam: PF00606; Glycoprotein_B/1.
 DR ProDom: PD000693; Glycoprot_B/1.
 KW Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 67
 FT CHAIN 1 67
 FT DOMAIN 68 932 GLYCOPROTEIN I.
 FT TRANSMEM 68 758 EXTRACELLULAR.
 FT DOMAIN 759 827 POTENTIAL.
 FT TRANSMEM 828 932 CYTOPLASMIC.
 FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 483 483 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 640 640 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 706 706 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 932 AA; 10195 MW; 9DCEA85C5FC3DA3 CRC64;

Query Match 37.9%; Score 44; DB 1; Length 932;
 Best Local Similarity 40.0%; Pred. No. 42;
 Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 REVYHPLATSHQGYF 15
 ||:|:|:|:|
 Db 659 RELVEPTANHKRYF 673

Search completed: July 12, 2004, 14:08:02
 Job time : 6.90698 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 14:01:53 ; Search time 11 Seconds
(without alignments)
192.383 Million cell updates/sec

Title: US-10-010-667a-2_COPY_92_113
Perfect score: 116
Sequence: 1 REVHPLATSHQQFYFKIPILV 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52.5	45.3	416	2 E64140	hypothetical prote
2	49	42.2	177	2 G84273	peptidic methionine
3	47.5	40.9	320	2 S73186	cytochrome f - red
4	47	40.5	233	2 D84797	hypothetical prote
5	47	40.5	431	2 H84392	O-acetyl homoserin
6	47	40.5	565	2 S52682	hypothetical prote
7	47	40.5	675	2 T39727	nucleoporin homolo
8	46	39.7	169	2 T07623	extensin homolo H
9	46	39.7	296	2 A11178	transcription regu
10	46	39.7	425	2 AD1149	O-acetylhomoserine
11	46	39.7	425	2 AD1508	O-acetylhomoserine
12	46	39.7	428	2 T08576	phenylalanine-cRNA
13	46	39.7	430	2 C83975	O-acetylhomoserine
14	46	39.7	735	2 AD0341	probable membrane
15	46	39.7	737	2 C70770	hypothetical prote
16	46	39.7	1839	1 OYBYK	adenylate cyclase
17	46	38.8	304	2 T09357	hypothetical prote
18	45	38.8	428	2 T46025	hypothetical prote
19	45	38.8	607	2 AG3323	penicillin-binding
20	44.5	38.4	136	2 S72508	sperm motility inh
21	44.5	38.4	1176	2 A27826	DNA-directed RNA p
22	44	37.9	163	2 C91092	hypothetical prote
23	44	37.9	163	2 G85937	hypothetical prote
24	44	37.9	472	2 G81293	probable type II p
25	44	37.9	771	2 S51421	hypothetical prote
26	44	37.9	928	1 VGBEBG	glycoprotein gi pr
27	44	37.9	932	1 VGBEBG	glycoprotein gi pr
28	44	37.9	969	2 B87336	isoleucyl-tRNA syn
29	43.5	37.5	326	2 S56534	hypothetical 36.9K

30	43.5	37.5	326	2 D91287	hypothetical prote
31	43.5	37.5	326	2 G86128	hypothetical prote
32	43	37.1	284	2 S71227	extensin 1 - Arabi
33	43	37.1	295	2 F83356	hypothetical prote
34	43	37.1	336	2 F86319	hypothetical prote
35	43	37.1	337	1 S14523	aspartate-semialde
36	43	37.1	338	2 B82118	probable aspartate
37	43	37.1	373	2 B96798	extensin (imported
38	43	37.1	379	2 C66873	transcription regu
39	43	37.1	384	2 B72324	oxaloacetate decar
40	43	37.1	385	2 T16447	hypothetical prote
41	43	37.1	426	2 A86634	O-acetylhomoserine
42	43	37.1	430	2 D72324	O-acetylhomoserine
43	43	37.1	434	2 AH3397	cysteine synthase
44	43	37.1	442	2 T44655	O-acetylhomoserine
45	43	37.1	485	2 T37550	hypothetical coile

ALIGNMENTS

RESULT 1
E64140
hypothetical protein H1021 - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 23-Sep-2002
C:Accession: E64140
R:Feilschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kierlavage, J.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhmann, J.L.; Geoghegan, N.S.M. Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Frazer, C.M.; Smith, H.O.; Venter A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: E64140
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1416 <RTGP>
A:Cross-references: GB:U32887; GB:I42023; NID:91572955; PIDN:AAC21699.1; PID:91572965;
A:Note: Best homolog was a hypothetical protein from Klebsiella pneumoniae
C:Genetics:
A:Start codon: GTG
C:Superfamily: citG protein

Query Match 45.3%; Score 52.5; DB 2; Length 416;
Best local Similarity 50.0%; Pred. No. 1.5;
Matches 10; Conservative 4; Mismatches 5; Indels 1; Gaps 1;

QY 1 REVHPLATSHQQFYFKIPILV 20
Db 34 KEITRPLETGHEAVFV-LPI 52

RESULT 2
G84273
peptide methionine sulfoxide reductase [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C:Accession: G84273
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, J.; Leibhauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jahn Jung, K.H.; Alam, M.; Freitas, T. Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: G84273
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-177 <STO>
A:Cross-references: GB:AE004437; NID:910580713; PIDN:AA619555.1; GSPDB:GN00138
C:Genetics:
A:Gene: msrA

RESULT 5

A:Cross-references: EMBL:AL109846; PIDN:CAB52802.1; GSPDB:GN00067; SPDB:SPBCT17G9 04C

A;Experimental source: strain 972h-; cosmid c17G9

C:Genetics:
A:Gene: SPDB:SPBC17G9.04c
A:Map position: 2

Query Match 40.5%; Score 47; DB 2; Length 675;
Best Local Similarity 57.1%; Pred. No. 20;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 VIHPLATSHQYFYK 16
Db 566 IISPAATHQYFYK 579

RESULT 8
T07623
extensin homolog HRGP2 - soybean (fragment)
N:Alternate names: hydroxyproline-rich glycoprotein HRGP2
C:Species: Glycine max (soybean)
C>Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 11-May-2000
C:Accession: T07623
R:Hong, J.C.; Cheong, Y.H.; Nagao, R.T.; Bahk, J.D.; Cho, M.J.; Key, J.L.
Plant Physiol. 104, 793-796, 1994
A>Title: Isolation and characterization of three soybean extensin cDNAs.
A:Reference number: Z16058; MUID:94211912; PMID:8159793
A:Accession: T07623
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-169 <HON>
A:Cross-references: EMBL:L22030; NID:g347454; PIDN:AAA33971.1; PID:g347455
A:Experimental source: strain Wayne; seedling
C:Genetics:
A:Gene: HRGP2
C:Superfamily: hydroxyproline-rich glycoprotein
C:Keywords: glycoprotein; hydroxyproline

Query Match 39.7%; Score 46; DB 2; Length 169;
Best Local Similarity 57.1%; Pred. No. 6;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 6 PLATSHQYFYKIP 19
Db 92 PSPTSHPRYYKSP 105

RESULT 9
A11178
transcription regulator homolog lmo0833 [imported] - *Listeria monocytogenes* (strain EGD-C:Species: *Listeria monocytogenes*
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: A11178
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, U.; Kuhn, M.; Kunst, F.; Kurapkát, G.; Madueno, E.; Maltournam, A.; Maok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlant, A.; Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: A11178
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-296 <GLA>
A:Cross-references: GB:NC_003210; PIDN:CAC98911.1; PID:g16410222; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo0833

Query Match 39.7%; Score 46; DB 2; Length 296;
Best Local Similarity 41.2%; Pred. No. 11;
Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Qy 2 EVIHPATSHQYFYKI 18
Db 376 IISPAATHQYFYKI 18

Db 222 EILHETSTHDCAFYRI 238

RESULT 10
AD1149

O-acetylhomoserine sulphydrylase homolog lmo0595 [imported] - *Listeria monocytogenes* (s
C:Species: *Listeria monocytogenes*
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AD1149
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, U.; Kuhn, M.; Kunst, F.; Kurapkát, G.; Madueno, E.; Maltournam, A.; Maok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlant, A.; Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AD1149
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-425 <GLA>
A:Cross-references: GB:NC_003210; PIDN:CAC98674.1; PID:g16409971; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo0595
C:Superfamily: O-succinylhomoserine (thiol)-lyase

Query Match 39.7%; Score 46; DB 2; Length 425;
Best Local Similarity 63.6%; Pred. No. 17;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VIHPLATSHQY 13
Db 376 IISPAATHQY 386

RESULT 11
AD1508
O-acetylhomoserine sulphydrylase homolog lmo0604 [imported] - *Listeria innocua* (strain
C:Species: *Listeria innocua*
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AD1508
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, U.; Kuhn, M.; Kunst, F.; Kurapkát, G.; Madueno, E.; Maltournam, A.; Maok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlant, A.; Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AD1508
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-425 <GLA>
A:Cross-references: GB:AL592022; PIDN:CAC95836.1; PID:g16413044; GSPDB:GN00178
A:Experimental source: strain C1p11262
C:Genetics:
A:Gene: lmo0604
C:Superfamily: O-succinylhomoserine (thiol)-lyase

Query Match 39.7%; Score 46; DB 2; Length 425;
Best Local Similarity 63.6%; Pred. No. 17;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VIHPLATSHQY 13
Db 376 IISPAATHQY 386

RESULT 12
T08576
phenylalanine-tRNA ligase (EC 6.1.1.20) beta chain T22F8.180 [similarity] - *Arabidopsis thaliana* (mouse-ear cress)
C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C/Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 03-Jun-2002
C/Accession: T08576
R/By: M.; Zimmermann, W.; Grueneisen, A.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; May
submitted to the Protein Sequence Database, May 1999
A/Reference number: Z16442
A/Accession: T08576
A/Molecule type: DNA
A/Residues: 1-428 <BRV>
A/Cross-references: EMBL:AL050351; GSPDB:GN00062; ATSP:T22F8.180
A/Experimental source: cultivar Columbia; BAC clone T22F8
C/Genetics:
A/Gene: ATSP:T22F8.180
A/Map position: 4
A/Intons: 49/3; 91/3; 121/3; 140/3; 179/2; 196/3; 213/1; 241/2; 273/3; 306/1; 330/1; 35
C/Superfamily: Yeast cytosolic phenylalanine-tRNA ligase, beta chain
C/Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 39.7%; Score 46; DB 2; Length 428;
Best Local Similarity 46.7%; Pred. No. 17;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 5 HPLATSHQOYFKIP 19
|||:|:|:|:
Db 200 HPARDSDHDFELKVP 214

RESULT 13

C83975
O-acetylhomoserine sulphydrylase BH2603 [imported] - Bacillus halodurans (strain C-125)
C/Species: Bacillus halodurans
C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C/Accession: C83975
R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A/Reference number: A83650; WUID:20512582; PMID:11058132
A/Accession: C83975
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-430 <STO>
A/Cross-references: GB:AP001516; GB:BA000004; NID:g10175192; PIDN:BA06322.1; GSPDB:GN00
A/Experimental source: strain C-125
C/Genetics:
A/Gene: BH2603
C/Superfamily: O-succinylhomoserine (thiol)-lyase

Query Match 39.7%; Score 46; DB 2; Length 430;
Best Local Similarity 63.6%; Pred. No. 17;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 VTHPLATSHQO 13
:|:|:|:|:|:
Db 379 ITHPASTHQO 389

RESULT 14

AD0341
probable membrane protein YPO2801 [imported] - Yersinia pestis (strain CO92)
C/Species: Yersinia pestis
C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C/Accession: AD0341
R/Parhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Parraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
Il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A/Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A/Reference number: AB0001; WUID:21470413; PMID:11586360
A/Accession: AD0341
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-735 <KUR>
A/Cross-references: GB:AL590842; PIDN:CA093035.1; PID:g15980773; GSPDB:GN00175
C/Genetics:

A/Gene: YPO2801

Query Match 39.7%; Score 46; DB 2; Length 735;
Best Local Similarity 50.0%; Pred. No. 32;
Matches 10; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 3 VTHPLATSHQOYFKIP 22
:|:|:|:|:|:
Db 529 LTHPLADPNOQLFFIP 548

RESULT 15

C70770
hypothetical protein RV1327c - Mycobacterium tuberculosis (strain H37Rv)
C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C/Accession: C70770
R/Cole, S.T.; Brosch, R.; Parhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
i Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: A70500; WUID:98295987; PMID:9634230
A/Accession: C70770
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-737 <COL>
A/Cross-references: GB:E273902; GB:AL123456; NID:g3261576; PIDN:CA098091.1; PID:e245021;
A/Experimental source: strain H37Rv
C/Genetics:
A/Gene: RV1327c

Query Match 39.7%; Score 46; DB 2; Length 737;
Best Local Similarity 46.7%; Pred. No. 32;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 5 HPLATSHQOYFKIP 19
|||:|:|:|:
Db 389 HPAREHRCQWFTLP 403

Search completed: July 12, 2004, 14:11:18
Job time : 12 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 14:10:33 ; Search time 38.8837 Seconds

(without alignments)
176.480 Million cell updates/sec

Title: US-10-010-667A-2_COPY_92_113
Perfect score: 116
Sequence: 1 REVHPLATSHQYFYKIPITV 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1279676 seqs, 311918243 residues

Total number of hits satisfying chosen parameters: 1279676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	116	100.0	22	14	US-10-011-095-19
2	116	100.0	22	14	US-10-010-667A-19
3	116	100.0	22	14	US-10-165-044-37
4	116	100.0	104	9	US-09-864-761-47234
5	116	100.0	267	9	US-09-747-835A-50
6	116	100.0	267	9	US-09-747-835A-51
7	116	100.0	267	12	US-10-312-312-50
8	116	100.0	267	12	US-10-312-312-51
9	116	100.0	339	9	US-09-759-143-879
10	116	100.0	339	9	US-09-780-669-879
11	116	100.0	339	9	US-09-822-827-879
12	116	100.0	339	9	US-09-802-520-11
13	116	100.0	339	9	US-09-895-793-879
14	116	100.0	339	9	US-09-895-814-879
15	116	100.0	339	12	US-10-205-267-13

16	116	100.0	339	12	US-10-408-009-2	Sequence 2, Appli
17	116	100.0	339	13	US-10-012-896-879	Sequence 879, App
18	116	100.0	339	14	US-10-011-095-2	Sequence 2, Appli
19	116	100.0	339	14	US-10-010-667A-2	Sequence 2, Appli
20	116	100.0	339	14	US-10-205-823-397	Sequence 397, App
21	116	100.0	339	14	US-10-144-678A-879	Sequence 879, App
22	116	100.0	339	14	US-10-294-025-879	Sequence 879, App
23	116	100.0	339	15	US-10-239-607-37	Sequence 879, App
24	116	100.0	339	15	US-10-295-027-714	Sequence 714, App
25	116	100.0	339	15	US-10-295-027-1347	Sequence 1347, App
26	116	100.0	368	12	US-10-425-114-72779	Sequence 72779, A
27	116	100.0	375	14	US-10-165-044-2	Sequence 2, Appli
28	68	58.6	26	14	US-10-165-044-40	Sequence 40, Appli
29	68	58.6	419	12	US-10-455-822-11	Sequence 11, Appli
30	68	58.6	419	12	US-10-455-822-80	Sequence 80, Appli
31	68	58.6	419	12	US-10-455-822-172	Sequence 172, Appli
32	68	58.6	419	12	US-10-455-822-174	Sequence 174, App
33	68	58.6	419	12	US-10-455-822-195	Sequence 195, App
34	68	58.6	419	15	US-10-239-607-32	Sequence 32, Appli
35	68	58.6	443	12	US-10-455-822-93	Sequence 93, Appli
36	68	58.6	444	12	US-10-455-822-88	Sequence 88, Appli
37	68	58.6	444	12	US-10-455-822-89	Sequence 89, Appli
38	68	58.6	444	12	US-10-455-822-90	Sequence 90, Appli
39	68	58.6	444	12	US-10-455-822-91	Sequence 91, Appli
40	68	58.6	444	12	US-10-455-822-92	Sequence 92, Appli
41	68	58.6	454	10	US-09-888-257A-10	Sequence 10, Appli
42	68	58.6	454	12	US-10-455-822-3	Sequence 3, Appli
43	68	58.6	454	12	US-10-455-822-7	Sequence 7, Appli
44	68	58.6	454	12	US-10-455-822-9	Sequence 9, Appli
45	68	58.6	454	12	US-10-455-822-19	Sequence 19, Appli

ALIGNMENTS

RESULT 1
US-10-011-095-19
; Sequence 19, Application US/10011095
; Publication No. US20030045682A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Hubert, Rene S.
; APPLICANT: Leong, Kahon
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Saffran, Douglas C.
; APPLICANT: Mitchell, Steve Chapell
; TITLE OF INVENTION: ANTIBODIES IMMUNOSPECIFIC FOR STEAP1 (AS AMENDED)
; FILE REFERENCE: 511582001610
; CURRENT APPLICATION NUMBER: US/10/011,095
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: 09/323,873
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: STEAP-1 peptide
US-10-011-095-19
Query Match 100.0%, Score 116, DB 14,
Best Local Similarity 100.0%, Pred. No. 3.4e-11,
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 REVHPLATSHQYFYKIPITV 22
Db 1 REVHPLATSHQYFYKIPITV 22

```
RESULT 2
US-10-010-667A-19
; Sequence 19, Application US/10010667A
; Publication No. US20030055217A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Hubert, Rene S.
; APPLICANT: Leong, Kahan
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Saffran, Douglas C.
; APPLICANT: Mitchell, Steve Chappell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; FILE REFERENCE: 511582001601
; CURRENT APPLICATION NUMBER: US/10/010,667A
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: 09/323,873
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 19
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: STEAP-1 peptide
US-10-010-667A-19

Query Match      100.0%; Score 116; DB 14; Length 22;
Best Local Similarity 100.0%; Pred. No. 3.4e-11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 REVHPLATSHQOYFYKIPILV 22
Db      1 REVHPLATSHQOYFYKIPILV 22

RESULT 3
US-10-165-044-37
; Sequence 37, Application US/10165044
; Publication No. US20030149531A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Rene S. Hubert
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas Saffran
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Steven Chappell Mitchell
; APPLICANT: Mary Faris
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; FILE REFERENCE: 51158-20016.02
; CURRENT APPLICATION NUMBER: US/10/165,044
; CURRENT FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: US 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: US 60/091,183
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 09/323,873
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: US 09/455,486
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: WO 99/62941
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/33040
; PRIOR FILING DATE: 2000-12-06
```

```
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 37
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-165-044-37

Query Match      100.0%; Score 116; DB 14; Length 22;
Best Local Similarity 100.0%; Pred. No. 3.4e-11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 REVHPLATSHQOYFYKIPILV 22
Db      1 REVHPLATSHQOYFYKIPILV 22

RESULT 4
US-09-864-761-47234
; Sequence 47234, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO: 47234
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```
/
/ FEATURE:
/ OTHER INFORMATION: MAP TO AC005053.1
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
/ OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.87
/ OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.57
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.59
/ OTHER INFORMATION: SWISSPROT HIT: O67305, EVALU6 5.60e+00
/ OTHER INFORMATION: EST_HUMAN HIT: BE875216.1, EVALU6 2.00e-54
US-09-864-761-47234

Query Match      100.0%; Score 116; DB 9; Length 104;
Best Local Similarity 100.0%; Pred. No. 1.8e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 REVTHPLATSHQQYFYKIPILV 22
        |||||||
Db      59 REVTHPLATSHQQYFYKIPILV 80

RESULT 5
US-09-747-835A-50
/ Sequence 50, Application US/09747835A
/ Patent No. US20020146692A1
/ GENERAL INFORMATION:
/ APPLICANT: Yamazaki, Victoria
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Zhou, Ping
/ APPLICANT: Wang, Dunrui
/ APPLICANT: Zhang, Jie
/ APPLICANT: Ren, Feiyan
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Dimanac, Radoje T
/ TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIKE
/ FILE REFERENCE: HVS-37CIP
/ CURRENT APPLICATION NUMBER: US/09/747,835A
/ PRIOR FILING DATE: 2002-03-08
/ PRIOR APPLICATION NUMBER: US 09/729,739
/ PRIOR FILING DATE: 2000-12-04
/ PRIOR APPLICATION NUMBER: US 09/653,450
/ PRIOR FILING DATE: 2000-08-31
/ PRIOR APPLICATION NUMBER: US 09/620,312
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: US 09/598,042
/ PRIOR FILING DATE: 2000-06-20
/ PRIOR APPLICATION NUMBER: US 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: US 09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 63
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 50
/ LENGTH: 267
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-747-835A-50

Query Match      100.0%; Score 116; DB 9; Length 267;
Best Local Similarity 100.0%; Pred. No. 5e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 REVTHPLATSHQQYFYKIPILV 22
        |||||||
Db      26 REVTHPLATSHQQYFYKIPILV 47

RESULT 6
US-09-747-835A-51
/ Sequence 51, Application US/09747835A
/ Patent No. US20020146692A1
/ GENERAL INFORMATION:
/ APPLICANT: Yamazaki, Victoria
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Zhou, Ping
/ APPLICANT: Wang, Dunrui
/ APPLICANT: Zhang, Jie
/ APPLICANT: Ren, Feiyan
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Dimanac, Radoje T
/ TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIKE
/ FILE REFERENCE: HVS-37CIP
/ CURRENT APPLICATION NUMBER: US/09/747,835A
/ PRIOR FILING DATE: 2002-03-08
/ PRIOR APPLICATION NUMBER: US 09/729,739
/ PRIOR FILING DATE: 2000-12-04
/ PRIOR APPLICATION NUMBER: US 09/653,450
/ PRIOR FILING DATE: 2000-08-31
/ PRIOR APPLICATION NUMBER: US 09/620,312
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: US 09/598,042
/ PRIOR FILING DATE: 2000-06-20
/ PRIOR APPLICATION NUMBER: US 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: US 09/488,725
```

```
/
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Zhou, Ping
/ APPLICANT: Wang, Dunrui
/ APPLICANT: Zhang, Jie
/ APPLICANT: Ren, Feiyan
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Dimanac, Radoje T
/ TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIKE
/ FILE REFERENCE: HVS-37CIP
/ CURRENT APPLICATION NUMBER: US/09/747,835A
/ PRIOR FILING DATE: 2002-03-08
/ PRIOR APPLICATION NUMBER: US 09/729,739
/ PRIOR FILING DATE: 2000-12-04
/ PRIOR APPLICATION NUMBER: US 09/653,450
/ PRIOR FILING DATE: 2000-08-31
/ PRIOR APPLICATION NUMBER: US 09/620,312
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: US 09/598,042
/ PRIOR FILING DATE: 2000-06-20
/ PRIOR APPLICATION NUMBER: US 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: US 09/488,725
US-09-747-835A-51

Query Match      100.0%; Score 116; DB 9; Length 267;
Best Local Similarity 100.0%; Pred. No. 5e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 REVTHPLATSHQQYFYKIPILV 22
        |||||||
Db      26 REVTHPLATSHQQYFYKIPILV 47

RESULT 7
US-10-312-312-50
/ Sequence 50, Application US/10312312
/ Publication No. US20040068097A1
/ GENERAL INFORMATION:
/ APPLICANT: Hyseq, Inc.
/ APPLICANT: Yamazaki, Victoria
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Zhou, Ping
/ APPLICANT: Wang, Dunrui
/ APPLICANT: Zhang, Jie
/ APPLICANT: Ren, Feiyan
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Dimanac, Radoje T
/ TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIKE
/ FILE REFERENCE: 21272-015-061/HVS-37CIP
/ CURRENT APPLICATION NUMBER: US/10/312,312
/ PRIOR FILING DATE: 2002-12-10
/ PRIOR APPLICATION NUMBER: US 09/729,739
/ PRIOR FILING DATE: 2000-12-04
/ PRIOR APPLICATION NUMBER: US 09/653,450
/ PRIOR FILING DATE: 2000-08-31
/ PRIOR APPLICATION NUMBER: US 09/620,312
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: US 09/598,042
/ PRIOR FILING DATE: 2000-06-20
/ PRIOR APPLICATION NUMBER: US 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: US 09/488,725
```

PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 63
SOFTWARE: PatentIn version 3.0
SEQ ID NO 50
LENGTH: 267
TYPE: PRT
ORGANISM: Homo sapiens
US-10-312-312-50

Query Match 100.0%; Score 116; DB 12; Length 267;
Best Local Similarity 100.0%; Pred. No. 5e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHQOYFYKIPILV 22
DB 26 REVHPLATSHQOYFYKIPILV 47

RESULT 8
US-10-312-312-51

Sequence 51, Application US/10312312
Publication NO. US20040068097A1
GENERAL INFORMATION:
APPLICANT: Hysq, Inc.
APPLICANT: Yamazaki, Victoria
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Zhou, Ping
APPLICANT: Wang, Dunrui
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyun
APPLICANT: Asundi, Vinod
APPLICANT: Drmanac, Radoje T
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIKE
FILE REFERENCE: 21272-015-061/HVS-37CIP
CURRENT FILING DATE: 2002-12-10
PRIOR FILING DATE: 2002-12-10
PRIOR FILING DATE: 2000-12-10
PRIOR FILING DATE: 2000-12-04
PRIOR FILING DATE: 2000-08-31
PRIOR FILING DATE: 2000-07-19
PRIOR FILING DATE: 2000-06-20
PRIOR FILING DATE: 2000-04-25
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 63
SOFTWARE: PatentIn version 3.0
SEQ ID NO 51
LENGTH: 267
TYPE: PRT
ORGANISM: Homo sapiens
US-10-312-312-51

Query Match 100.0%; Score 116; DB 12; Length 267;
Best Local Similarity 100.0%; Pred. No. 5e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHQOYFYKIPILV 22
DB 26 REVHPLATSHQOYFYKIPILV 47

RESULT 9
US-09-759-143-879

Sequence 879, Application US/09759143
Patent No. US2002022248A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiaqichun

APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yuqi
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
APPLICANT: Fanger, Gary R.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darriek
APPLICANT: Li, Samuel
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.427C23
CURRENT FILING DATE: 2001-01-12
CURRENT FILING DATE: 2001-01-12
NUMBER OF SEQ ID NOS: 934
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 879
LENGTH: 339
TYPE: PRT
ORGANISM: Homo sapiens
US-09-759-143-879

Query Match 100.0%; Score 116; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 6.5e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHQOYFYKIPILV 22
DB 92 REVHPLATSHQOYFYKIPILV 113

RESULT 10
US-09-780-669-879

Sequence 879, Application US/09780669
Patent No. US2002051977A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiaqichun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yuqi
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
APPLICANT: Fanger, Gary R.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darriek
APPLICANT: Li, Samuel
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William
APPLICANT: Hural, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.427C24
CURRENT FILING DATE: 2001-02-09
NUMBER OF SEQ ID NOS: 943
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 879
LENGTH: 339
TYPE: PRT

; ORGANISM: Homo sapiens
US-09-780-669-879

Query Match 100.0%; Score 116; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 6.5e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHQOYFYKIPILV 22
|||||
DB 92 REVHPLATSHQOYFYKIPILV 113

RESULT 11

US-09-822-827-879
; Sequence 879, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C1
; CURRENT APPLICATION NUMBER: US/09/822,827
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-822-827-879

Query Match 100.0%; Score 116; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 6.5e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHQOYFYKIPILV 22
|||||
DB 92 REVHPLATSHQOYFYKIPILV 113

RESULT 12
US-09-802-520-11
; Sequence 11, Application US/09802520
; Publication No. US20020187472A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Faris, Mary
; APPLICANT: Chen, Huel-Mei
; APPLICANT: Ison, Craig H.
; TITLE OF INVENTION: STEAP-RELATED PROTEIN
; FILE REFERENCE: PC-0037 US
; CURRENT APPLICATION NUMBER: US/09/802,520
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PERL Program
; SEQ ID NO 11
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020187472A1 96572948
US-09-802-520-11

Query Match 100.0%; Score 116; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 6.5e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHQOYFYKIPILV 22
|||||
DB 92 REVHPLATSHQOYFYKIPILV 113

RESULT 13

US-09-895-793-879
; Sequence 879, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darriek
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C2
; CURRENT APPLICATION NUMBER: US/09/895,793
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-895-793-879

Query Match 100.0%; Score 116; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 6.5e-10;
Matches 22; Conservative 100.0%; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHQOYFYKIPILV 22
|||||
DB 92 REVHPLATSHQOYFYKIPILV 113

RESULT 14
US-09-895-814-879
; Sequence 879, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darriek
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.

```

; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Ranger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C26
; CURRENT APPLICATION NUMBER: US/09/895,814
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-895-814-879

```

```

Query Match          100.0%; Score 116; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 6.5e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 REVHPLATSHQOYFYKIPILV 22
      |||||
Db      92 REVHPLATSHQOYFYKIPILV 113

```

```

RESULT 15
US-10-205-267-13
; Sequence 13, Application US/10205267
; Publication No. US20030064397A1
; GENERAL INFORMATION:
; APPLICANT: Spence, Kimberly M.
; APPLICANT: Rickett Paula K.
; APPLICANT: Lal, Preeti G.
; APPLICANT: Ison, Craig H.
; TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN PROSTATE AND LU
; TITLE OF INVENTION: TUMORS
; FILE REFERENCE: PV-0008 CIP
; CURRENT APPLICATION NUMBER: US/10/205,267
; CURRENT FILING DATE: 2002-07-24
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PERL Program
; SEQ ID NO 13
; LENGTH: 339
; TYPE: PRT
; ORGANISM:
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: Genbank ID No: g6572948
US-10-205-267-13

```

```

Query Match          100.0%; Score 116; DB 12; Length 339;
Best Local Similarity 100.0%; Pred. No. 6.5e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 REVHPLATSHQOYFYKIPILV 22
      |||||
Db      92 REVHPLATSHQOYFYKIPILV 113

```

Search completed: July 12, 2004, 14:27:49
 Job time : 39.8837 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 14:04:08 ; Search time 13.5581 Seconds
(without alignments)
83.770 Million cell updates/sec

Title: US-10-010-667a-2_COPY_92_113
Perfect score: 116
Sequence: 1 REVHPLATSHQYFYKIPILIV 22

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCITUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	116	100.0	22	4	US-09-323-873A-19
2	116	100.0	339	4	US-09-323-873A-2
3	116	100.0	339	4	US-09-685-166A-879
4	46	37.9	195	4	US-09-134-001C-3548
5	44	37.1	933	2	US-08-682-847-2
6	43	37.1	337	4	US-09-252-931A-25798
7	42.5	36.6	245	4	US-09-489-039A-8972
8	42	36.2	220	4	US-09-134-000C-4003
9	42	36.2	284	3	US-09-357-251-24
10	42	36.2	319	4	US-09-134-001C-3688
11	42	36.2	1140	4	US-08-471-112A-4
12	42	36.2	1809	3	US-09-012-515A-12
13	42	36.2	1809	3	US-08-360-144A-12
14	42	36.2	1809	4	US-09-012-504A-12
15	42	36.2	1809	4	US-09-012-399A-12
16	42	36.2	2549	4	US-08-471-112A-3
17	42	36.2	2549	4	US-08-265-967C-1
18	42	36.2	2549	4	US-08-305-790B-2
19	42	36.2	2549	5	PCR-US95-06722-12
20	41.5	35.8	625	3	US-08-581-148C-18
21	41.5	35.8	625	4	US-08-759-436-3
22	41.5	35.8	625	4	US-08-759-436-5
23	41	35.3	159	4	US-09-252-991A-18982
24	41	35.3	190	1	US-08-106-981-2
25	41	35.3	322	4	US-09-540-236-2884
26	41	35.3	344	4	US-09-489-039A-8542
27	41	35.3	392	2	US-08-282-197C-60

28	41	35.3	396	3	US-09-461-474-2	Sequence 2, Appli
29	41	35.3	592	2	US-08-366-490-2	Sequence 2, Appli
30	41	35.3	592	3	US-08-860-483A-2	Sequence 2, Appli
31	41	35.3	879	1	US-08-220-151-2	Sequence 2, Appli
32	41	35.3	879	1	US-08-220-151-3	Sequence 3, Appli
33	41	35.3	879	1	US-08-413-118-2	Sequence 2, Appli
34	41	35.3	879	1	US-08-413-118-3	Sequence 3, Appli
35	41	35.3	879	1	US-08-413-118-106	Sequence 106, App
36	41	35.3	879	3	US-08-473-446-2	Sequence 2, Appli
37	41	35.3	879	3	US-08-473-446-3	Sequence 3, Appli
38	41	35.3	879	3	US-08-473-446-106	Sequence 106, App
39	40.5	34.9	261	4	US-09-107-532A-4856	Sequence 4856, Ap
40	40	34.5	68	4	US-09-134-001C-2858	Sequence 2858, Ap
41	40	34.5	77	4	US-09-328-352-7432	Sequence 7432, Ap
42	40	34.5	115	2	US-07-903-029-6	Sequence 6, Appli
43	40	34.5	206	4	US-09-219-797-2	Sequence 2, Appli
44	40	34.5	303	1	US-08-185-432-5	Sequence 5, Appli
45	40	34.5	306	4	US-09-107-532A-5464	Sequence 5464, Ap

ALIGNMENTS

```
RESULT 1
US-09-323-873A-19
; Sequence 19, Application US/09323873A
; Patent No. 6329503
; GENERAL INFORMATION:
; APPLICANT: Daniel R. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Kahan Leong
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Steve Chappell Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; FILE REFERENCE: 129.16USU2
; CURRENT APPLICATION NUMBER: US/09/323,873A
; CURRENT FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: STRAP-1 PEPTIDE
; US-09-323-873A-19

Query Match          100.0%; Score 116; DB 4; Length 22;
Best Local Similarity 100.0%; Pred. No. 6.7e-12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 REVHPLATSHQYFYKIPILIV 22
Db       1 REVHPLATSHQYFYKIPILIV 22

RESULT 2
US-09-323-873A-2
; Sequence 2, Application US/09323873A
; Patent No. 6329503
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Kahan Leong
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Steve Chappell Mitchell
```

;; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
;; TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USES THEREOF
;; FILE REFERENCE: 129 16US02
;; CURRENT APPLICATION NUMBER: US/09/323,873A
;; CURRENT FILING DATE: 1999-06-01
;; PRIOR APPLICATION NUMBER: 60/087,520
;; PRIOR FILING DATE: 1998-06-01
;; PRIOR APPLICATION NUMBER: 60/091,183
;; PRIOR FILING DATE: 1998-06-30
;; NUMBER OF SEQ ID NOS: 32
;; SOFTWARE: FASTSEQ for Windows Version 4.0
;; SEQ ID NO 2
;; LENGTH: 339
;; TYPE: PRT
;; ORGANISM: Homo Sapiens
US-09-323-873A-2

Query Match 100.0%; Score 116; DB 4; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHOOYFYKIPILV 22
DB 92 REVHPLATSHOOYFYKIPILV 113

RESULT 3
US-09-685-166A-879
; Sequence 879, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, David C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan U.
; APPLICANT: Wang, Yugu
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasser A.W.
; APPLICANT: Hegler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685,166A
; CURRENT FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-685-166A-879

Query Match 100.0%; Score 116; DB 4; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHOOYFYKIPILV 22
DB 92 REVHPLATSHOOYFYKIPILV 113

RESULT 4
US-09-134-001C-3548
; Sequence 3548, Application US/09134001C

;; Patent No. 6380370
;; GENERAL INFORMATION:
;; APPLICANT: Lynn Doucette-Stamm et al
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
;; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;; FILE REFERENCE: GTC-007
;; CURRENT APPLICATION NUMBER: US/09/134,001C
;; CURRENT FILING DATE: 1998-08-13
;; PRIOR APPLICATION NUMBER: US 60/064,964
;; PRIOR FILING DATE: 1997-11-08
;; PRIOR APPLICATION NUMBER: US 60/055,779
;; PRIOR FILING DATE: 1997-08-14
;; NUMBER OF SEQ ID NOS: 5674
;; SEQ ID NO 3548
;; LENGTH: 195
;; TYPE: PRT
;; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3548

Query Match 39.7%; Score 46; DB 4; Length 195;
Best Local Similarity 44.4%; Pred. No. 9.2;
Matches 8; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 1 REVHPLATSHOOYFYKI 18
DB 128 RSLSPKTSYQEMOYQI 145

RESULT 5
US-08-682-847-2
; Sequence 2, Application US/08682847
; Patent No. 5858989
; GENERAL INFORMATION:
; APPLICANT: BABLUK, LORNE
; APPLICANT: VAN DEN HURK, SYLVIA
; APPLICANT: ZAMB, TIM
; APPLICANT: FITZPATRICK, DAVID
; TITLE OF INVENTION: RECOMBINANT BOVINE HERPESVIRUS TYPE 1
; TITLE OF INVENTION: POLYPEPTIDES AND VACCINES
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & ROESTER
; STREET: 755 PAGE MILL ROAD
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/682,847
; FILING DATE: 12-JUL-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: PARK, FREDIE K.
; REGISTRATION NUMBER: 35,636
; REFERENCE/DOCKET NUMBER: 29310-20005.10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 933 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-682-847-2

Query Match 37.9%; Score 44; DB 2; Length 933;

Best Local Similarity 40.0%; Pred. No. 1,1e+02;
Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 REVHPLATSHQOYF 15
||:|:|:|:|:|
Db 660 RELVEPCANMKRYF 674

RESULT 6
US-09-252-991A-25798

; Sequence 25798, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 25798

; LENGTH: 337

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-25798

Query Match 37.1%; Score 43; DB 4; Length 337;
Best Local Similarity 46.2%; Pred. No. 51;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 REVHPLATSHQO 13
||:|:|:|:|:|
Db 240 REVHPIVTHHE 252

RESULT 7

US-09-489-039A-8972

; Sequence 8972, Application US/09489039A

; Patent No. 6610836

; GENERAL INFORMATION:

; APPLICANT: Gary Bretton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

; FILE REFERENCE: 2709.2004001

; CURRENT APPLICATION NUMBER: US/09/489,039A

; PRIOR FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: US 60/117,747

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 8972

; LENGTH: 245

; TYPE: PRT

; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-8972

Query Match 36.6%; Score 42.5; DB 4; Length 245;
Best Local Similarity 35.7%; Pred. No. 43;
Matches 10; Conservative 3; Mismatches 6; Indels 9; Gaps 1;

Qy 1 REVHPLATSHQOYF 19
||:|:|:|:|:|
Db 201 RHITTEIATAPKPFYVADDDHQYLYKNP 228

RESULT 8

US-09-134-000C-4003

; Sequence 4003, Application US/09134000C

; Patent No. 6617156

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

; FILE REFERENCE: 032796-032

; CURRENT APPLICATION NUMBER: US/09/134,000C

; PRIOR FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/055,778

; PRIOR FILING DATE: 1997-08-15

; NUMBER OF SEQ ID NOS: 6812

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 4003

; LENGTH: 220

; TYPE: PRT

; ORGANISM: Enterococcus faecalis

US-09-134-000C-4003

Query Match 36.2%; Score 42; DB 4; Length 220;
Best Local Similarity 41.2%; Pred. No. 45;
Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 3 VIHPLATSHQOYFYKIP 19
||:|:|:|:|:|
Db 153 IVRVIMTKYQOEFYFEP 169

RESULT 9

US-09-357-251-24

; Sequence 24, Application US/09357251

; Patent No. 6271441

; GENERAL INFORMATION:

; APPLICANT: Falco, S. Carl

; APPLICANT: Farnold, Lavo O.

; APPLICANT: Orozco, Buddy

; APPLICANT: Schwab, James S.

; TITLE OF INVENTION: Plant Aminoacyl-tRNA Synthetase

; FILE REFERENCE: BB-1193

; CURRENT APPLICATION NUMBER: US/09/357,251

; PRIOR FILING DATE: 1999-07-20

; EARLIER APPLICATION NUMBER: 60/093,530

; EARLIER FILING DATE: July 21, 1998

; NUMBER OF SEQ ID NOS: 37

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 24

; LENGTH: 284

; TYPE: PRT

; ORGANISM: Triticum aestivum

US-09-357-251-24

Query Match 36.2%; Score 42; DB 3; Length 284;
Best Local Similarity 46.7%; Pred. No. 60;
Matches 7; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 5 HPLATSHQOYFYKIP 19
||:|:|:|:|:|
Db 55 HPARDSDHTEFLKAP 69

RESULT 10

US-09-134-001C-3888

; Sequence 3888, Application US/09134001C

; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C

; PRIOR FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; NUMBER OF SEQ ID NOS: 5674

SEQ ID NO 3888
LENGTH: 319
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3888

Query Match 36.2%; Score 42; DB 4; Length 319;
Best Local Similarity 47.1%; Pred. No. 69;
Matches 8; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 4 IHPLATSHOOQYFYKIP 20
DB 249 IHKXCSDHOORFYNCIS 265

RESULT 11
US-08-471-112A-4
Sequence 4, Application US/08471112A
Patent No. 6313264
GENERAL INFORMATION:
APPLICANT: Molnar-Kimber, Katherine L.
APPLICANT: Falli, Amedeo F.
APPLICANT: Caggiano, Thomas J.
APPLICANT: Nakanishi, Koji
APPLICANT: Chen, Yangu
TITLE OF INVENTION: EFFECTOR PROTEINS OF RAPAMYCIN
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner, L.L.P.
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,112A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/384,524
FILING DATE: 13-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/312,023
FILING DATE: 26-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/207,975
FILING DATE: 08-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Siekman, Michael T.
REGISTRATION NUMBER: 36,276
REFERENCE/DOCKET NUMBER: 01142.0058-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1140 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-471-112A-4

Query Match 36.2%; Score 42; DB 4; Length 1140;
Best Local Similarity 50.0%; Pred. No. 2.8e+02;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 REVYHPLATSHOOQYFY 16
DB 274 ROLDHPLPTVHPQVTV 289

RESULT 12
US-09-012-515A-12
Sequence 12, Application US/09012515A
Patent No. 6127521
GENERAL INFORMATION:
APPLICANT: Berlin, Vivian
APPLICANT: Chiu, Maria Isabel
APPLICANT: Cottarel, Guillaume
APPLICANT: Damagnez, Veronique
TITLE OF INVENTION: IMMUNOSUPPRESSANT TARGET PROTEINS
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/012,515A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/360,144
FILING DATE: 20-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: APV-036.02
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1809 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-012-515A-12

Query Match 36.2%; Score 42; DB 3; Length 1809;
Best Local Similarity 50.0%; Pred. No. 4.7e+02;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 REVYHPLATSHOOQYFY 16
DB 943 ROLDHPLPTVHPQVTV 958

RESULT 13
US-08-360-144A-12
Sequence 12, Application US/08360144A
Patent No. 6150137
GENERAL INFORMATION:
APPLICANT: Berlin, Vivian
APPLICANT: Chiu, Maria Isabel
APPLICANT: Cottarel, Guillaume
APPLICANT: Damagnez, Veronique
TITLE OF INVENTION: IMMUNOSUPPRESSANT TARGET PROTEINS
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square

CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/360.144A
FILING DATE: 20-DEC-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: APV-036.02
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1809 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-360-144A-12

Query Match 36.2%; Score 42; DB 4; Length 1809;
Best Local Similarity 50.0%; Pred. No. 4.7e+02;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 1 REVHPLATSHOQYFY 16
Db 943 RQDHPLEPTVHPQVY 958

RESULT 14
US-09-012-504A-12
Sequence 12, Application US/09012504A
GENERAL INFORMATION:
APPLICANT: Berlin, V.
APPLICANT: Chiu, I.
APPLICANT: Cottarel, G.
APPLICANT: Damagnez, V.
TITLE OF INVENTION: IMMUNOSUPPRESSANT TARGET PROTEINS
FILE REFERENCE: APRI-P05-036
CURRENT APPLICATION NUMBER: US/09/012.504A
CURRENT FILING DATE: 1998-01-23
PRIOR APPLICATION NUMBER: 08/360.144
PRIOR FILING DATE: 1994-12-20
PRIOR APPLICATION NUMBER: 08/250.795
PRIOR FILING DATE: 1994-05-27
NUMBER OF SEQ ID NOS: 35
SOFTWARE: Patentin version 3.1
SEQ ID NO 12
LENGTH: 1809
TYPE: PRT
ORGANISM: Mammalian
US-09-012-504A-12

Query Match 36.2%; Score 42; DB 4; Length 1809;
Best Local Similarity 50.0%; Pred. No. 4.7e+02;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 1 REVHPLATSHOQYFY 16
Db 943 RQDHPLEPTVHPQVY 958

RESULT 15
US-09-012-399A-12

Sequence 12, Application US/09012399A
Patent No. 6509152
GENERAL INFORMATION:
APPLICANT: Berlin, Vivian
APPLICANT: Chiu, Maria Isabel
APPLICANT: Cottarel, Guillaume
APPLICANT: Damagnez, Veronique
TITLE OF INVENTION: IMMUNOSUPPRESSANT TARGET PROTEINS
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/012.399A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/360.144
FILING DATE: 20-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: APV-036.02
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1809 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-012-399A-12

Query Match 36.2%; Score 42; DB 4; Length 1809;
Best Local Similarity 50.0%; Pred. No. 4.7e+02;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 1 REVHPLATSHOQYFY 16
Db 943 RQDHPLEPTVHPQVY 958

Search completed: July 12, 2004, 14:12:23
Job time : 13.5581 secs

This Page Blank (uspto)

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 13:51:07 ; Search time 48.3488 Seconds
(without alignments)
128.567 Million cell updates/sec

Title: US-10-010-667a-2_COPY_92_113
Perfect score: 116
Sequence: 1 REVHPLATSHQYFYKIPILV 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: A_Geneseqp29Jan04:*
 - 2: geneseqp1990s:*
 - 3: geneseqp2000s:*
 - 4: geneseqp2001s:*
 - 5: geneseqp2002s:*
 - 6: geneseqp2003as:*
 - 7: geneseqp2003bs:*
 - 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	116	100.0	22	3	AA58198 Human STR
2	116	100.0	22	4	AAE02786 Extracellular
3	116	100.0	104	4	AB40417 Peptide #
4	116	100.0	104	4	AAW73944 Aam73944 Human Don
5	116	100.0	104	4	ABG55696 Human liv
6	116	100.0	104	5	ABG43832 Human pep
7	116	100.0	254	6	ABU98426 STEAP-1 v
8	116	100.0	254	6	ABU98425 STEAP-1 v
9	116	100.0	255	6	ABU98429 STEAP-1 v
10	116	100.0	255	6	ABU98428 STEAP-1 v
11	116	100.0	258	6	ABU98395 Novel hum
12	116	100.0	258	6	ABU98389 Novel hum
13	116	100.0	258	6	ABU98424 STEAP-1 v
14	116	100.0	258	6	ABU98392 Novel hum
15	116	100.0	258	6	ABU98387 Novel hum
16	116	100.0	258	6	ABU98393 Novel hum
17	116	100.0	258	6	ABU98390 Novel hum
18	116	100.0	258	6	ABU98391 Novel hum
19	116	100.0	258	6	ABU98384 Novel hum
20	116	100.0	258	6	ABU98431 STEAP-1 v
21	116	100.0	258	6	ABU98396 Novel hum
22	116	100.0	258	6	ABU98397 Novel hum
23	116	100.0	258	6	ABU98394 Novel hum
24	116	100.0	258	6	ABU98388 Novel hum
25	116	100.0	258	6	ABU98386 Novel hum

26	116	100.0	258	6	ABU98399 Novel hum
27	116	100.0	258	6	ABU98398 Novel hum
28	116	100.0	267	6	ABU60886 Human G P
29	116	100.0	267	6	ABU60887 Human G P
30	116	100.0	282	6	ABU98432 STEAP-1 v
31	116	100.0	282	6	ABU98427 STEAP-1 v
32	116	100.0	282	6	ABU98385 Novel hum
33	116	100.0	339	3	AA58194 Human STR
34	116	100.0	339	4	AAW01282 P789P ant
35	116	100.0	339	4	AAU69927 Human pro
36	116	100.0	339	4	AAW78845 Human pro
37	116	100.0	339	4	ABU71818 Prostate
38	116	100.0	339	5	ABR95387 Human P78
39	116	100.0	339	5	ABG61813 Prostate
40	116	100.0	339	6	ABU98383 Novel hum
41	116	100.0	339	6	ABU98414 STEAP-1 v
42	116	100.0	339	6	ABU98430 STEAP-1 v
43	116	100.0	339	6	ABR54499 Prostate
44	116	100.0	339	7	ABU63313 Human six
45	116	100.0	339	7	ABD75573 Prostate

ALIGNMENTS

RESULT 1
AA58198
ID AA58198 standard; peptide; 22 AA.
XX
AC AA58198;
XX
DT 14-MAR-2000 (first entry)
XX
DE Human STRAP-1 peptide, corresponding to STRAP-1 extracellular region 1.
XX
KW Serpentine transmembrane antigen of the prostate; STRAP-1; prostate;
KW transmembrane domain; type IIIa membrane protein; expression; cancer;
KW prostate cancer; bladder cancer; colon cancer; pancreatic cancer;
KW ovarian cancer; tumour antigen; immunisation; immune response; cellular;
KW humoral; anticancer vaccine; antibody; detection; diagnosis; prognosis;
KW monitoring; susceptibility; therapeutic inhibitor; drug targeting;
KW recombinant protein.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9962941-A2.
XX
PD 09-DEC-1999.
XX
PF 01-JUN-1999; 99WO-US012157.
XX
PR 01-JUN-1998; 98US-0087520P.
XX 30-JUN-1998; 98US-0091183P.
XX (UROG-) UROGENESYS INC.
PA (AFAR/) AFAR D E.
PA (HUBE/) HUBERT R S.
PA (LEON/) LEONG K.
PA (RAIT/) RAITANO A B.
PA (SAFF/) SAFFRAN D C.
PI Afar DE, Hubert RS, Leong K, Raitano AB, Saffran DC;
XX WPI; 2000-072832/06.
XX Novel proteins useful as diagnostic markers and therapeutic targets,
XX particularly for prostatic cancer.
XX Disclosure; Page 22; 83pp; English.
XX Sequences AA58198-Y58200 represent synthetic peptides that correspond to
CC the extracellular regions of STRAP-1 (serpentine transmembrane antigen of

the prostate, AAY58194). These peptides were used to raise monoclonal anti-STRAP-1 antibodies. STRAP-1 is the prototype member of the STRAP family of proteins (AAY58194-Y58197) which exhibit a high degree of structural conservation, but which show no significant structural homology to known human proteins. The STRAP-1 gene has been localised to chromosome 7p22. STRAP-1 is thought to be a type IIIa membrane protein and is expressed predominantly in prostate cells in normal human tissues. Structurally, STRAP-1 is a 339 amino acid protein characterised by six transmembrane domains and intracellular N- and C-termini, suggesting that it folds in a "serpentine" manner into three extracellular and two intracellular loops. STRAP-1 mRNA and protein expression is maintained at high levels and throughout all stages of prostate cancer. STRAP-1 mRNA and/or protein is also overexpressed in certain other cancers, including bladder, colon, pancreatic and ovarian cancer. The function of the STRAP proteins is not known. They may be ion channels (from the presence of six transmembrane domains, a feature which is shared by certain ion channels) or gap-junction proteins (from immunohistochemical staining). STRAP-1 and STRAP-2 are cell-surface tumour antigens. Immunisation with a STRAP protein induces cellular and humoral immune responses against STRAP-expressing cells. STRAP proteins may be used to identify specific-binding agents, to produce anticancer vaccines and to generate specific antibodies. The antibodies may be used for detection, prognosis, and monitoring of cancers (or susceptibility to cancer), as therapeutic inhibitors or to target therapeutic agents to their site of action. STRAP nucleic acids may be used for recombinant protein production, as diagnostic and prognostic reagents, for identifying STRAP-expressing cells for screening inhibitors of STRAP expression and for therapeutic modulation/inhibition of STRAP expression. Since high levels of STRAP proteins are exposed on the cell surface, and because they are easily targeted by systemically administered agents, and because they are expressed mainly on prostatic epithelial cells, agents targeted to them should have minimal side effects on other tissues

Sequence 22 AA:

Query Match 100.0%; Score 116; DB 3; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.4e-11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHQOYFYKPIIV 22
DB 1 REVHPLATSHQOYFYKPIIV 22

RESULT 2

AAE02786
ID AAE02786 standard; peptide; 22 AA.

AC AAE02786;

DT 06-AUG-2001 (first entry)

DE Extracellular loop #1 of human STRAP-1, suitable for cloning into pFc.

XX Human; cytosolic; antiproliferative; vaccine; gene therapy;

KW six transmembrane epithelial antigen of the prostate-1; STRAP-1;

KM chromosome 7p22.3; cancer; prostate; colon; bladder; pancreatic; lung;

KW ovarian; extracellular loop; serpentine transmembrane antigen.

XX Homo sapiens.

OS WO200140276-A2.

PN 07-JUN-2001.

PD 06-DEC-2000; 2000WO-US033040.

PR 06-DEC-1999; 99US-00455486.

XX (UROG-) UROGENESYS INC.

PA Afar DEH, Hubert RS, Raitano AB, Saffran DC, Mitchell SC;
PI Faris M, Jakobovits A;

XX WPI, 2001-367804/38.

DR New STEAP (six transmembrane epithelial antigen of the prostate) PT proteins, expressed in human cancers, useful for detecting and treating cancer.

XX Example 19, Page 102; 187pp; English.

XX The present invention relates to human six transmembrane epithelial CC antigen of the prostate (STEAP) protein. STEAP is a member of cell CC surface serpentine transmembrane antigens. STEAP gene is used in gene CC therapy. Inhibiting the development or progression of a cancer (eg. CC prostate, colon, bladder, lung, ovarian and pancreatic) expressing STEAP CC or inhibiting growth or killing cells expressing STEAP in a patient, CC comprises administering a vaccine composition to the patient. Treating a CC patient with a cancer that expresses STEAP, or inhibiting growth or CC killing cells expressing STEAP, comprises administering to the patient a CC vector encoding single chain monoclonal antibody that comprises the CC variable domains of the heavy and light chains of the monoclonal antibody CC that specifically binds to STEAP, such that the vector delivers the CC single chain monoclonal antibody coding sequence to the cancer cells and CC the encoded single chain monoclonal antibody is expressed CC intracellularly. The present sequence is extracellular loop of STEAP-1 CC suitable for cloning into pFc, which is used in the invention. STEAP-1 CC gene is located on chromosome 7p22.3

Sequence 22 AA:

Query Match 100.0%; Score 116; DB 4; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.4e-11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHQOYFYKPIIV 22
DB 1 REVHPLATSHQOYFYKPIIV 22

RESULT 3

ABB40417
ID ABB40417 standard; peptide; 104 AA.

AC ABB40417;

DT 04-FEB-2002 (first entry)

DE Peptide #7923 encoded by human foetal liver single exon probe.

XX Human; foetal liver; gene expression; single exon nucleic acid probe.

OS Homo sapiens.

PN WO200157277-A2.

PD 09-AUG-2001.

PR 30-JAN-2001; 2001WO-US000669.

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI, 2001-483447/52.

Human genome-derived single exon nucleic acid probes useful for analyzing

PT gene expression in human fetal liver.
XX
PS Claim 27; SEQ ID NO 33052; 639bp + Sequence Listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human fetal liver. The
CC present sequence is a peptide encoded by a single exon nucleic acid probe
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 104 AA;

Query Match 100.0%; Score 116; DB 4; Length 104;
Best Local Similarity 100.0%; Pred. No. 7.4e-11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHQGYFYKIPILV 22
Db 59 REVHPLATSHQGYFYKIPILV 80

RESULT 4
AAM73944
ID AAM73944 standard; protein; 104 AA.
XX
AC AAM73944;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 34250.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001MO-US000664.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human bone marrow.
XX
PS Example 4; SEQ ID NO 34250; 658bp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention
XX
SQ Sequence 104 AA;

Query Match 100.0%; Score 116; DB 4; Length 104;
Best Local Similarity 100.0%; Pred. No. 7.4e-11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHQGYFYKIPILV 22
Db 59 REVHPLATSHQGYFYKIPILV 80

RESULT 5
ABG55696
ID ABG55696 standard; peptide; 104 AA.
XX
AC ABG55696;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human liver peptide, SEQ ID NO 34344.
XX
KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
KW hypercholesterolaemia; coronary heart disease.
XX
OS Homo sapiens.
XX
PN WO200157273-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001MO-US000664.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human adult liver.
XX
PS Claim 27; SEQ ID NO 34344; 658bp; English.
XX
CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (II) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABG47348-ABG59930 represent human
CC liver single exon encoded peptides of the invention. Note: The sequence
CC information for this patent does not appear in the printed specification
CC but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 104 AA;

Query Match 100.0%; Score 116; DB 4; Length 104;
Best Local Similarity 100.0%; Pred. No. 7.4e-11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHQGYFYKIPILV 22
Db 59 REVHPLATSHQGYFYKIPILV 80

RESULT 6
ABG43832
ID ABG43832 standard; peptide; 104 AA.
XX
AC ABG43832;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human peptide encoded by genome-derived single exon probe SEQ ID 33497.
XX
XX Human; single exon probe; asthma; lung cancer; COPD; IID;
XX chronic obstructive pulmonary disease; interstitial lung disease;
XX familial idiopathic pulmonary fibrosis; neurofibromatosis;
XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
XX Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
XX pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
XX pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
XX primary ciliary dyskinesia; pulmonary hypertension;
XX hyaline membrane disease.
XX
OS Homo sapiens.
XX
PN W0200186003-A2.
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US000665.
XX
PR 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2002-114183/15.
XX
PT Spatially-addressable set of single exon nucleic acid probes, used to
XX measure gene expression in human lung samples.
XX
PS Claim 27; SEQ ID NO 33497; 634BP; English.
XX
XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human lung comprising single exon nucleic acid probes having one of
XX 12614 nucleic acid sequences mentioned in the specification, or their
XX complements or the 12387 open reading frames derived from the 12614
XX probes. Also included are a microarray comprising the novel set of probes
XX; the novel set of probes which hybridize at high stringency to a nucleic
XX acid expressed in the human lung; measuring gene expression in a sample
XX derived from human lung; comprising (a) contacting the array with a
XX collection of detectably labeled nucleic acids derived from human lung
XX mRNA, and (b) measuring the label detectably bound to each probe of the
XX array; identifying exons in a eukaryotic genome, comprising (a)
XX algorithmically predicting at least one exon from genomic sequences of
XX the eukaryote; and (b) detecting specific hybridisation of detectably
XX labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
XX having a fragment identical to the predicted exon, the probe is included
XX in the above mentioned microarray; assigning exons to a single gene,
XX comprising (a) identifying exons from genomic sequence by the method
XX above and (b) measuring the expression of each of the exons in several
XX tissues and/or cell types using hybridisation to a single exon
XX microarrays having a probe with the exon, where a common pattern of
XX expression of the exons in the tissues and/or cell types indicates that
XX the exons should be assigned to a single gene; a peptide comprising one
XX of 12011 sequences, mentioned in the specification, or encoded by the

CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (IID), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a peptide/protein encoded by a single exon probe of
CC the invention. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 104 AA;
XX
Query Match 100.0%; Score 116; DB 5; Length 104;
Best Local Similarity 100.0%; Pred. No. 7,4e-11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 REVHPLATSHOOYFYKIPILV 22
DB 59 REVHPLATSHOOYFYKIPILV 80
XX
RESULT 7
ABU98426
ID ABU98426 standard; protein; 254 AA.
XX
AC ABU98426;
XX
DT 31-JUL-2003 (first entry)
XX
DE STEAP-1 variant 8PID4 v.2 #2.
XX
XX STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
XX cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
XX vaccine.
XX
XX Homo sapiens.
XX
XX W02003022995-A2.
XX
XX 20-MAR-2003.
XX
XX 06-SEP-2002; 2002WO-US028371.
XX
XX 06-SEP-2001; 2001US-0317840P.
XX 05-APR-2002; 2002US-0370387P.
XX
XX (AGEN-) AGENSYS INC.
XX
XX Paris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;
XX
XX WPI; 2003-313240/30.
XX
XX New composition comprising a substance that modulates the status of a
XX STEAP-1-related protein, useful for treating and detecting cancer.
XX
XX Example 53; Page 169-170; 248BP; English.
XX
XX The invention describes a composition comprising a substance that
XX modulates the status of a protein (I) of 340 or 283 amino acids, or of
XX any of the 15 sequences of 259 amino acids, given in the specification,
XX or a molecule that is modulated by the protein, where the status of the
XX cell that expresses the protein is modulated. The compositions, proteins,
XX polynucleotides and methods are useful for treating and detecting cancer.
XX The STEAP-1-related proteins are useful for generating cancer vaccines.
XX The polynucleotides are useful as tools for delineating, with greater
XX precision, cytogenetic abnormalities in the chromosomal region that
XX encodes STEAP-1 that may contribute to the malignant phenotype. This is
XX the amino acid sequence of a variant of human six transmembrane

CC epithelial antigen of the prostate or STEAP-1
XX
SQ Sequence 254 AA;

Query Match 100.0%; Score 116; DB 6; Length 254;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHQGYFYKIPILV 22
|||||
Db 92 REVHPLATSHQGYFYKIPILV 113

RESULT 8
ABU98425
ID ABU98425 standard; protein; 254 AA.

XX ABU98425;

AC 31-JUL-2003 (first entry)

DT STEAP-1 variant 8P1D4 v.1 #1.

XX STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
KW cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
XX vaccine.

OS Homo sapiens.

XX WO2003022995-A2.

PN 20-MAR-2003.

PD 06-SEP-2002; 2002WO-US028371.

PF 06-SEP-2001; 2001US-0317840P.

PR 05-APR-2002; 2002US-0370387P.

XX (AGEN-) AGENSYS INC.

PA Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;

PI WPI; 2003-313240/30.

DR New composition comprising a substance that modulates the status of a

PT STEAP-1-related protein, useful for treating and detecting cancer.

XX Example 53; Page 169-170; 248pp; English.

CC The invention describes a composition comprising a substance that
modulates the status of a protein (I) of 340 or 283 amino acids, or of
any of the 15 sequences of 259 amino acids, given in the specification,
or a molecule that is modulated by the protein, where the status of the
cell that expresses the protein is modulated. The compositions, proteins,
polynucleotides and methods are useful for treating and detecting cancer.
CC The STEAP-1-related proteins are useful for generating cancer vaccines.
CC The polynucleotides are useful as tools for delineating, with greater
precision, cytogenetic abnormalities in the chromosomal region that
encodes STEAP-1 that may contribute to the malignant phenotype. This is
the amino acid sequence of a variant of human six transmembrane
epithelial antigen of the prostate or STEAP-1

CC Sequence 254 AA;

SQ

Query Match 100.0%; Score 116; DB 6; Length 254;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHQGYFYKIPILV 22
|||||
Db 92 REVHPLATSHQGYFYKIPILV 113

RESULT 9
ABU98429
ID ABU98429 standard; protein; 255 AA.

XX ABU98429;

AC 31-JUL-2003 (first entry)

DT STEAP-1 variant 8P1D4 v.3 #2.

XX STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
KW cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
XX vaccine.

OS Homo sapiens.

XX WO2003022995-A2.

PN 20-MAR-2003.

PD 06-SEP-2002; 2002WO-US028371.

PF 06-SEP-2001; 2001US-0317840P.

PR 05-APR-2002; 2002US-0370387P.

XX (AGEN-) AGENSYS INC.

PA Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;

PI WPI; 2003-313240/30.

DR New composition comprising a substance that modulates the status of a

PT STEAP-1-related protein, useful for treating and detecting cancer.

XX Example 53; Page 172; 248pp; English.

CC The invention describes a composition comprising a substance that
modulates the status of a protein (I) of 340 or 283 amino acids, or of
any of the 15 sequences of 259 amino acids, given in the specification,
or a molecule that is modulated by the protein, where the status of the
cell that expresses the protein is modulated. The compositions, proteins,
polynucleotides and methods are useful for treating and detecting cancer.
CC The STEAP-1-related proteins are useful for generating cancer vaccines.
CC The polynucleotides are useful as tools for delineating, with greater
precision, cytogenetic abnormalities in the chromosomal region that
encodes STEAP-1 that may contribute to the malignant phenotype. This is
the amino acid sequence of a variant of human six transmembrane
epithelial antigen of the prostate or STEAP-1

CC Sequence 255 AA;

Query Match 100.0%; Score 116; DB 6; Length 255;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHQGYFYKIPILV 22
|||||
Db 92 REVHPLATSHQGYFYKIPILV 113

RESULT 10

ABU98428
ID ABU98428 standard; protein; 255 AA.

XX ABU98428;

AC 31-JUL-2003 (first entry)

DT STEAP-1 variant 8P1D4 v.2 #3.

XX STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
KW cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
XX vaccine.

```
XX OS Homo sapiens.
XX XX
XX PN WO2003022995-A2.
XX PD
XX PF 20-MAR-2003.
XX PR 06-SEP-2002; 2002WO-US028371.
XX PR 06-SEP-2001; 2001US-0317840P.
XX PR 05-APR-2002; 2002US-0370387P.
XX PA (AGEN-) AGENSYS INC.
XX PI Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;
XX WPI; 2003-313240/30.
XX DR
XX PT New composition comprising a substance that modulates the status of a
XX PT STEAP-1-related protein, useful for treating and detecting cancer.
XX PS Example 53; Page 172; 248pp; English.
XX CC The invention describes a composition comprising a substance that
XX CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
XX CC any of the 15 sequences of 259 amino acids, given in the specification,
XX CC or a molecule that is modulated by the protein, where the status of the
XX CC cell that expresses the protein is modulated. The compositions, proteins,
XX CC polynucleotides and methods are useful for treating and detecting cancer.
XX CC The STEAP-1-related proteins are useful for generating cancer vaccines.
XX CC The polynucleotides are useful as tools for delineating, with greater
XX CC precision, cytogenetic abnormalities in the chromosomal region that
XX CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
XX CC the amino acid sequence of a variant of human six transmembrane
XX CC epithelial antigen of the prostate or STEAP-1
XX SQ Sequence 255 AA;

Query Match 100.0%; Score 116; DB 6; Length 255;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHQYFYKIPILV 22
Db |||||
92 REVHPLATSHQYFYKIPILV 113

RESULT 11
ABU98395
ID ABU98395 standard; protein; 258 AA.
XX AC
XX ABU98395;
XX DT 31-JUL-2003 (first entry)
XX DE Novel human gene STEAP-1 variant 13.
XX XX
XX STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
XX KM cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
XX KM vaccine.
XX OS Homo sapiens.
XX XX
XX PN WO2003022995-A2.
XX PD
XX PF 20-MAR-2003.
XX PR 06-SEP-2002; 2002WO-US028371.
XX PR 06-SEP-2001; 2001US-0317840P.
XX PR 05-APR-2002; 2002US-0370387P.
XX PA (AGEN-) AGENSYS INC.
```

```
XX PI Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;
XX XX
XX DR WPI; 2003-313240/30.
XX DR N-PSDB; ACD02609.
XX PT New composition comprising a substance that modulates the status of a
XX PT STEAP-1-related protein, useful for treating and detecting cancer.
XX PS Example 2; Fig 2M; 248pp; English.
XX CC The invention describes a composition comprising a substance that
XX CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
XX CC any of the 15 sequences of 259 amino acids, given in the specification,
XX CC or a molecule that is modulated by the protein, where the status of the
XX CC cell that expresses the protein is modulated. The compositions, proteins,
XX CC polynucleotides and methods are useful for treating and detecting cancer.
XX CC The STEAP-1-related proteins are useful for generating cancer vaccines.
XX CC The polynucleotides are useful as tools for delineating, with greater
XX CC precision, cytogenetic abnormalities in the chromosomal region that
XX CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
XX CC the amino acid sequence of a variant of human six transmembrane
XX CC epithelial antigen of the prostate or STEAP-1
XX SQ Sequence 258 AA;

Query Match 100.0%; Score 116; DB 6; Length 258;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHQYFYKIPILV 22
Db |||||
92 REVHPLATSHQYFYKIPILV 113

RESULT 12
ABU98389
ID ABU98389 standard; protein; 258 AA.
XX AC
XX ABU98389;
XX DT 31-JUL-2003 (first entry)
XX DE Novel human gene STEAP-1 variant 7.
XX XX
XX STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
XX KM cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
XX KM vaccine.
XX OS Homo sapiens.
XX XX
XX PN WO2003022995-A2.
XX PD
XX PF 20-MAR-2003.
XX PR 06-SEP-2002; 2002WO-US028371.
XX PR 06-SEP-2001; 2001US-0317840P.
XX PR 05-APR-2002; 2002US-0370387P.
XX PA (AGEN-) AGENSYS INC.
XX XX
XX PI Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;
XX XX
XX DR WPI; 2003-313240/30.
XX DR N-PSDB; ACD02609.
XX PT New composition comprising a substance that modulates the status of a
XX PT STEAP-1-related protein, useful for treating and detecting cancer.
XX PS Example 2; Fig 2G; 248pp; English.
XX CC The invention describes a composition comprising a substance that
```

modulates the status of a protein (1) of 340 or 283 amino acids, or of any of the 15 sequences of 259 amino acids, given in the specification, or a molecule that is modulated by the protein, where the status of the cell that expresses the protein is modulated. The compositions, proteins, polynucleotides and methods are useful for treating and detecting cancer. The STAP-1-related proteins are useful for generating cancer vaccines. The polynucleotides are useful as tools for delineating, with greater precision, cytogenetic abnormalities in the chromosomal region that encodes STAP-1 that may contribute to the malignant phenotype. This is the amino acid sequence of a variant of human six transmembrane epithelial antigen of the prostate or STAP-1

SQ Sequence 258 AA;

Query Match 100.0%; Score 116; DB 6; Length 258;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHQGYFYKIPILV 22
|||
Db 92 REVHPLATSHQGYFYKIPILV 113

RESULT 13

ABU98424 ABU98424 standard; protein; 258 AA.

XX AC ABU98424;

XX DT 31-JUL-2003 (first entry)

XX DE STAP-1 variant 8PID4 V.2 #1.

XX KM STAP-1; six transmembrane epithelial antigen of the prostate; cancer;
XX KW cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
XX KV vaccine.

XX OS Homo sapiens.

XX PN WO2003022995-A2.

XX PD 20-MAR-2003.

XX PF 06-SEP-2002; 2002WO-US028371.

XX PR 06-SEP-2001; 2001US-0317840P.

XX PR 05-APR-2002; 2002US-0370387P.

XX PA (AGEN-) AGENSYS INC.

XX PI Faris M, Ge W, Raitano AB, Chalilta-Eid PM, Jakobovits A;

XX DR WPI; 2003-313240/30.

XX PT New composition comprising a substance that modulates the status of a
XX PT STAP-1-related protein, useful for treating and detecting cancer.

XX PS Example 53; Page 169; 248pp; English.

XX CC The invention describes a composition comprising a substance that
XX CC modulates the status of a protein (1) of 340 or 283 amino acids, or of
XX CC any of the 15 sequences of 259 amino acids, given in the specification,
XX CC or a molecule that is modulated by the protein, where the status of the
XX CC cell that expresses the protein is modulated. The compositions, proteins,
XX CC polynucleotides and methods are useful for treating and detecting cancer.
XX CC The STAP-1-related proteins are useful for generating cancer vaccines.
XX CC The polynucleotides are useful as tools for delineating, with greater
XX CC precision, cytogenetic abnormalities in the chromosomal region that
XX CC encodes STAP-1 that may contribute to the malignant phenotype. This is
XX CC the amino acid sequence of a variant of human six transmembrane
XX CC epithelial antigen of the prostate or STAP-1

SQ Sequence 258 AA;

Query Match 100.0%; Score 116; DB 6; Length 258;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHQGYFYKIPILV 22
|||
Db 92 REVHPLATSHQGYFYKIPILV 113

RESULT 14

ABU98392 ABU98392 standard; protein; 258 AA.

XX AC ABU98392;

XX DT 31-JUL-2003 (first entry)

XX DE Novel human gene STAP-1 variant 10.

XX KM STAP-1; six transmembrane epithelial antigen of the prostate; cancer;
XX KW cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
XX KV vaccine.

XX OS Homo sapiens.

XX PN WO2003022995-A2.

XX PD 20-MAR-2003.

XX PF 06-SEP-2002; 2002WO-US028371.

XX PR 06-SEP-2001; 2001US-0317840P.

XX PR 05-APR-2002; 2002US-0370387P.

XX PA (AGEN-) AGENSYS INC.

XX PI Faris M, Ge W, Raitano AB, Chalilta-Eid PM, Jakobovits A;

XX DR WPI; 2003-313240/30.

XX DR N-PSDB; ACDD02606.

XX PT New composition comprising a substance that modulates the status of a
XX PT STAP-1-related protein, useful for treating and detecting cancer.

XX PS Example 2; Fig 2J; 248pp; English.

XX CC The invention describes a composition comprising a substance that
XX CC modulates the status of a protein (1) of 340 or 283 amino acids, or of
XX CC any of the 15 sequences of 259 amino acids, given in the specification,
XX CC or a molecule that is modulated by the protein, where the status of the
XX CC cell that expresses the protein is modulated. The compositions, proteins,
XX CC polynucleotides and methods are useful for treating and detecting cancer.
XX CC The STAP-1-related proteins are useful for generating cancer vaccines.
XX CC The polynucleotides are useful as tools for delineating, with greater
XX CC precision, cytogenetic abnormalities in the chromosomal region that
XX CC encodes STAP-1 that may contribute to the malignant phenotype. This is
XX CC the amino acid sequence of a variant of human six transmembrane
XX CC epithelial antigen of the prostate or STAP-1

SQ Sequence 258 AA;

Query Match 100.0%; Score 116; DB 6; Length 258;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHQGYFYKIPILV 22
|||
Db 92 REVHPLATSHQGYFYKIPILV 113

RESULT 15

ABU98387

ID ABU98387 standard; protein; 258 AA.
XX
AC ABU98387;
XX
DT 31-JUL-2003 (first entry)
XX
DE Novel human gene STEAP-1 variant 5.
XX
KW STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
KW cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
KW vaccine.
XX
OS Homo sapiens.
XX
PN WO2003022995-A2.
XX
PD 20-MAR-2003.
XX
PF 06-SEP-2002; 2002WO-US028371.
XX
PR 06-SEP-2001; 2001US-0317840P.
PR 05-APR-2002; 2002US-0370387P.
XX
PA (AGEN-) AGENSYS INC.
XX
PI Paris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;
XX
DR WPI; 2003-313240/30.
DR N-PSDB; ACD02601.
XX
PT New composition comprising a substance that modulates the status of a
PT STEAP-1-related protein, useful for treating and detecting cancer.
XX
PS Example 2; Fig 2E; 248pp; English.
XX
CC The invention describes a composition comprising a substance that
CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
CC any of the 15 sequences of 259 amino acids, given in the specification,
CC or a molecule that is modulated by the protein, where the status of the
CC cell that expresses the protein is modulated. The compositions, proteins,
CC polynucleotides and methods are useful for treating and detecting cancer.
CC The STEAP-1-related proteins are useful for generating cancer vaccines.
CC The polynucleotides are useful as tools for delineating, with greater
CC precision, cytogenetic abnormalities in the chromosomal region that
CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
CC the amino acid sequence of a variant of human six transmembrane
CC epithelial antigen of the prostate or STEAP-1
XX
SQ Sequence 258 AA;

Query Match 100.0%; Score 116; DB 6; Length 258;
Best Local Similarity 100.0%; Pred. No. 2e-10; Mismatches 0; Indels 0; Gaps 0;
Matches 22; Conservative 0;

OY 1 REVHPLATSHQOYFYKIPILV 22
|||
DB 92 REVHPLATSHQOYFYKIPILV 113

Search completed: July 12, 2004, 14:07:22
Job time : 48.3488 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 14:00:58 ; Search time 22.1512 Seconds

(Without alignments)
213.658 Million cell updates/sec

Title: US-10-010-667A-2_COPY_277_291

Perfect score: 93

Sequence: 1 WIDIKQFWYTPPTF 15

Scoring table: BLOSUM62

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database : SPTREMBL_25:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_virus:*

16: sp_bacteriopl:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	91	97.8	338	6	Q9GL50
2	80	86.0	339	11	Q924J9
3	80	86.0	339	11	Q924Z2
4	80	86.0	339	11	Q9CWR7
5	55	59.1	419	16	Q8F0P8
6	49	52.7	448	16	Q98CK3
7	48	51.6	246	15	Q8UHF2
8	48	51.6	453	16	Q8UHF8
9	46	49.5	122	15	Q9IUL9
10	46	49.5	844	15	Q97002
11	45	48.4	107	2	Q8EVC0
12	45	48.4	107	9	Q8EACJ
13	45	48.4	122	15	Q9UDN6
14	45	48.4	122	15	Q9QIY7
15	45	48.4	122	15	Q9YXR4
16	45	48.4	122	15	Q9YXQ1

17	45	48.4	122	15	Q9ILL6	Q9ILL6 human immun
18	45	48.4	122	15	Q9QIT8	Q9QIT8 human immun
19	45	48.4	122	15	Q7ZUR8	Q7ZUR8 human immun
20	45	48.4	122	15	Q7ZUR5	Q7ZUR5 human immun
21	45	48.4	133	15	Q8UQZ0	Q8UQZ0 human immun
22	45	48.4	133	15	Q8UQZ9	Q8UQZ9 human immun
23	45	48.4	133	15	Q8UR20	Q8UR20 human immun
24	45	48.4	151	15	Q7SM24	Q7SM24 human immun
25	45	48.4	153	15	Q7SM42	Q7SM42 human immun
26	45	48.4	155	15	Q8U3Q3	Q8U3Q3 human immun
27	45	48.4	161	15	Q8JER1	Q8JER1 human immun
28	45	48.4	163	15	Q9OE72	Q9OE72 human immun
29	45	48.4	167	15	Q8JER2	Q8JER2 human immun
30	45	48.4	169	15	Q8JER0	Q8JER0 human immun
31	45	48.4	169	15	Q8JER9	Q8JER9 human immun
32	45	48.4	214	15	Q8AU20	Q8AU20 human immun
33	45	48.4	358	15	Q78120	Q78120 human immun
34	45	48.4	375	11	Q8BUD3	Q8BUD3 mus musculus
35	45	48.4	483	16	Q8Z289	Q8Z289 enterococcus
36	45	48.4	711	15	Q7SVK4	Q7SVK4 human immun
37	45	48.4	711	15	Q7SVK3	Q7SVK3 human immun
38	45	48.4	713	15	Q7SVK2	Q7SVK2 human immun
39	45	48.4	790	15	Q8Q2X3	Q8Q2X3 human immun
40	45	48.4	833	15	Q9QK14	Q9QK14 human immun
41	45	48.4	833	15	Q9QKH9	Q9QKH9 human immun
42	45	48.4	833	15	Q9QK13	Q9QK13 human immun
43	45	48.4	840	15	Q8Q2X4	Q8Q2X4 human immun
44	45	48.4	841	15	Q9IV31	Q9IV31 human immun
45	45	48.4	846	15	Q89292	Q89292 human immun

ALIGNMENTS

RESULT 1

ID Q9GL50 PRELIMINARY: PRT: 338 AA.

AC Q9GL50: 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)

DE Six transmembrane endothelial antigen of PNEC.

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

OC NCBI_TaxID=9823;

RN [1]

RP SEQUENCE FROM N.A.

RA Nagasaka T., Boulday G., Coupel S., Coulon F., Tesson L., Heslan J.-M., Soullion J.-P., Chaireau B., "Differential gene expression in endothelial cells during TNF-alpha-RT and iPS-mediated activation."

RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF319659; AAG33868.1; -.

GO GO:0016021; C:integral to membrane; IEA.

KW Transmembrane.

SQ SEQUENCE 338 AA: 39918 MW: ED490E8E067A32B CRC64;

Query Match 97.8%; Score 91; DB 6; Length 338;
Best Local Similarity 86.7%; Pred. No. 1,le-06;
Matches 13; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WIDIKQFWYTPPTF 15
Db 276 WIDIKQFWYTPPTF 290

RESULT 2
ID Q924J9 PRELIMINARY: PRT: 339 AA.
AC Q924J9: 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

```
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Six-transmembrane epithelial antigen of the prostate.
GN STEAP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=21371909; PubMed=11479226;
RA Yang D., Holt G.E., Velchers M.P., Kwon E.D., Kast W.M.;
RT "Murine six-transmembrane epithelial antigen of the prostate, prostate
RT stem cell antigen, and prostate-specific membrane antigen: prostate-
RT specific cell-surface antigens highly expressed in prostate cancer of
RT transgenic adenocarcinoma mouse prostate mice.";
RL Cancer Res. 61:5857-5860(2001).
DR EMBL; AF297098; AAK83126.1; -.
DR MGD; MGI:1917608; Steap.
DR GO; GO:0016021; C:integral to membrane; IEA.
KW Transmembrane.
SQ SEQUENCE 339 AA; 39318 MW; 4B26A71FF559E84F CRC64;

Query Match
Best Local Similarity 86.0%; Score 80; DB 11; Length 339;
Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 WIDIKQFWYPTPTF 15
Db 277 WVDVSQFWYMPPTF 291

RESULT 3
Q924Z2 PRELIMINARY; PRT; 339 AA.
AC Q924Z2;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Dudalim.
GN STEAP OR 1010001D01RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Seru V., Manivet P., Lambdin D., Vaubourdoille M., Kellermann O.,
RA Loric S.;
RT "Prostate and non-prostate expression of dudalim, the mouse ortholog
RT of human STEAP.";
RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY029584; AAK50537.1; -.
DR MGD; MGI:1917608; Steap.
SQ SEQUENCE 339 AA; 39109 MW; 32A2C29F2E333BD0 CRC64;

Query Match
Best Local Similarity 86.0%; Score 80; DB 11; Length 339;
Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 WIDIKQFWYPTPTF 15
Db 277 WVDVSQFWYMPPTF 291

RESULT 4
Q9CWR7 PRELIMINARY; PRT; 339 AA.
AC Q9CWR7;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE 2410007B19RIK protein.
```

```
GN STEAP OR 2410007B19RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; Tissue=Embryonic stem cells;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai U., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aikawa T., Hara A., Fukunishi Y., Konno H., Adachi U., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schmitt L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake U., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli U., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyszewski A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK010437; BAB26938.1; -.
DR MGD; MGI:1917608; Steap.
SQ SEQUENCE 339 AA; 39264 MW; 3F7AB9C7520F0968 CRC64;

Query Match
Best Local Similarity 86.0%; Score 80; DB 11; Length 339;
Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 WIDIKQFWYPTPTF 15
Db 277 WVDVSQFWYMPPTF 291

RESULT 5
Q8F0P8 PRELIMINARY; PRT; 419 AA.
AC Q8F0P8;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Conserved hypothetical protein.
GN LA3444.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=1173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RT Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AB011502; AAN50642.1; -.
DR InterPro; IPR001466; Beta_lactamase.
DR InterPro; IPR000437; Prok_lipo_prot_S.
DR Pfam; PF00144; beta-lactamase_1.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 419 AA; 47652 MW; BE6FE7CF180F0B80 CRC64;

Query Match
Best Local Similarity 59.1%; Score 55; DB 16; Length 419;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 WIDIKQFWYPTPTF 14
```

Db 248 WIDFKDYEWYSPS 261

RESULT 6

Q98CK3 PRELIMINARY; PRT; 448 AA.

AC Q98CK3; 08JF92;

DT 01-OCT-2001 (TREMBLrel. 18, Created)

DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)

DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)

DE Alpha-glucosides ABC transporter, substrate-binding protein. ML513.

GN Rhizobium loti (Mesorhizobium loti).

OS Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Phyllobacteriaceae; Mesorhizobium.

CC NCBI_TaxID=381;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MAF030309;

RX MEDLINE=21082930; PubMed=11214968;

RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S., Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T., Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;

RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium Mesorhizobium loti."

RL DNA Res. 7:331-338(2000).

DR EMBL; AP003005; BAB51618.1; -.

DR GO; GO:0005215; P:transporter activity; IEA.

DR GO; GO:0006810; P:transport; IEA.

DR InterPro; IPR006059; SBP_bac_1.

DR Pfam; PF01547; SBP_bac_1; 1.

KW Complete proteome.

SQ SEQUENCE 448 AA; 48920 MW; 9B7B8569A6945D19 CRC64;

Query Match 52.7%; Score 49; DB 16; Length 448;

Best Local Similarity 53.8%; Pred. No. 10;

Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 3 DIKQFWYTPPTF 15

Db 164 DVKSLWYVSPDNF 176

RESULT 7

Q8JF92 PRELIMINARY; PRT; 246 AA.

AC Q8JF92;

DT 01-OCT-2002 (TREMBLrel. 22, Created)

DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)

DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)

DE Truncated envelope protein (Fragment).

OS Human immunodeficiency virus 1.

OC Viruses; Retroviridae; Lentivirus.

CC NCBI_TaxID=11676;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=RF;

RX MEDLINE=22011541; PubMed=12015905;

RA Jones D.R., Suzuki K., Pillier S.C.;

RT "A 100-Amino Acid Truncation in the Cytoplasmic Tail of Glycoprotein 41 in the Reference HIV Type 1 Strain RF."

RL AIDS Res. Hum. Retroviruses 18:513-517(2002).

DR EMBL; AF403705; AAM21476.1; -.

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0019031; C:lipid envelope; IEA.

DR GO; GO:0005198; F:structural molecule activity; IEA.

DR InterPro; IPR000328; Env_GP41.

DR Pfam; PF00517; GP41; 1.

KW Envelope protein; Transmembrane.

FT NON_TER 1

SQ SEQUENCE 246 AA; 27734 MW; 4F0058F96DE1C4B3 CRC64;

Query Match 51.6%; Score 48; DB 15; Length 246;

Best Local Similarity 70.0%; Pred. No. 8;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 WIDIKQFWY 10

Db 162 WEDITQWYV 171

RESULT 8

Q8JHT8 PRELIMINARY; PRT; 453 AA.

AC Q8JHT8;

DT 01-JUN-2002 (TREMBLrel. 21, Created)

DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)

DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)

DE ABC transporter, substrate binding protein.

GN AGLE OR AFU0591 OR AGR_C1045.

OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.

CC NCBI_TaxID=176299;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=21608550; PubMed=11743193;

RA Wood D.W., Sebubal J.C., Kaul R., Monks D.E., Kitajima J.P., Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L., Chen Y., Paulsen I.T., Eissen J.A., Karp P.D., Bovee D. Sr., Chapman P., Clendinning J., Deatherage G., Gillet M., Grant C., Kuyavlin T., Levy R., Li M.-J., McClelland E., Palmieri A., Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero F., Gordon D., Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M., Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M., Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V., Nester E.W.;

RT "The genome of the natural genetic engineer Agrobacterium tumefaciens C58."

RL Science 294:2317-2323(2001).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=21608551; PubMed=11743194;

RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M., Qurollo B., Goldman B.S., Gao Y., Askenazi M., Hailing C., Mullin L., Houtmel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F., Wollam C., Allinger M., Dougherty D., Scott C., Lapps C., Markelz B., Flanagan C., Crowell C., Garson J., Lomo C., Sear C., Strub G., Cielo C., Slater S.;

RT "Genome sequence of the plant pathogen and biotechnology agent Agrobacterium tumefaciens C58."

RL Science 294:2323-2328(2001).

DR EMBL; AE009027; AAL11608.1; -.

DR EMBL; AE007993; AAK86402.1; -.

DR PIR; A97431; A97431.

DR PIR; AB2649; AB2649.

DR GO; GO:0005215; F:transporter activity; IEA.

DR GO; GO:0006810; P:transport; IEA.

DR InterPro; IPR006059; SBP_bac_1.

DR Pfam; PF01547; SBP_bac_1; 1.

KW Complete proteome.

SQ SEQUENCE 453 AA; 49090 MW; 90896249313CD85E CRC64;

Query Match 51.6%; Score 48; DB 16; Length 453;

Best Local Similarity 53.8%; Pred. No. 15;

Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 3 DIKQFWYTPPTF 15

Db 169 DVKSLWYVSPDNF 181

RESULT 9

Q9IUI9

ID Q91JL9 PRELIMINARY; PRT; 122 AA.
AC Q91JL9;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AR3;
RX MEDLINE=20346416; PubMed=10890362;
RA Masciotra S., Liellera B., Belloso W., Clara L., Tanuri A., Ramos A.,
Bassig J., Lal R., Pieniazek D.;
RT "Evidence for a high frequency of HIV-1 subtype F infections among
RT heterosexual population in Buenos Aires, Argentina.";
RL AIDS Res. Hum. Retroviruses 16:1007-1014(2000).
DR EMBL: AF220713; AAF76832.1; -
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0019031; C:viral envelope; IEA.
DR InterPro: IPR000328; F:structural molecule activity; IEA.
DR Pfam: PF00517; GP41; 1. _GP41.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 122
SQ SEQUENCE 122 AA; 14840 MW; 10821A3BC39F4E6 CRC64;

Query Match
Best Local Similarity 49.5%; Score 46; DB 15; Length 122;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 WIDIKOFVWY 10
DB 112 WEDISQWLWY 121

RESULT 10
Q97002 PRELIMINARY; PRT; 844 AA.
AC Q97002;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HIVBR020.17;
RX MEDLINE=96190564; PubMed=8627686;
RA Gao F., Morrison S.G., Robertson D.L., Thornton C.L., Craig S.,
Karlsson G., Sodroski J., Morgado M., Galvao-Castro B.,
von Briesen H., Beddows S., Weber J., Sharp P.M., Shaw G.M.,
Hahn B.H.;
RT "Molecular cloning and analysis of functional envelope genes from
RT human immunodeficiency virus type 1 sequence subtypes A through G. The
RT WHO and NIAID Networks for HIV Isolation and Characterization.";
RL J. Virol. 70:1651-1657(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=HIVBR020.17;
RA Allan E.E.;
RT Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: U27401; AAB06242.1; -
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0019031; C:viral envelope; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.

DR InterPro: IPR000328; Env GP41.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 844 AA; 95594 MW; 0B02E379CEFD17691 CRC64;

Query Match
Best Local Similarity 49.5%; Score 46; DB 15; Length 844;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 WIDIKOFVWY 10
DB 660 WEDISQWLWY 669

RESULT 11
Q9EV00 PRELIMINARY; PRT; 107 AA.
AC Q9EV00;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E32511;
RX MEDLINE=20407286; PubMed=10948097;
RA Unkneif A., Schmidt H.;
RT "Structural analysis of phage-borne stx genes and their flanking
RT sequences in shiga toxin-producing escherichia coli and shiga11a
RT dysenteriae type 1 strains.";
RL Infect. Immun. 68:4856-4864(2000).
DR EMBL: AJ251452; CAC05542.1; -
KW Hypothetical protein.
SQ SEQUENCE 107 AA; 12177 MW; 8BCD6C52B7D29C66 CRC64;

Query Match
Best Local Similarity 48.7%; Score 45; DB 2; Length 107;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 WIDIKOFVWYTPPTF 15
DB 22 WCNIKILIMYOSDTF 36

RESULT 12
Q8HAJ1 PRELIMINARY; PRT; 107 AA.
AC Q8HAJ1;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein.
OS Bacteriophage ICI59.
OC Viruses.
OX NCBI_TaxID=210928;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=159;
RA Muniesa M., Jofre J.;
RT "Variability of shiga converting bacteriophages in E. coli O157:H7
RT strains of human origin isolated from the same outbreak.";
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF548456; AAN59922.1; -
KW Hypothetical protein.
SQ SEQUENCE 107 AA; 12197 MW; 8BCD6C52B7D29C66 CRC64;

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 13:52:02 ; Search time 4.7093 Seconds

(without alignments)
165.853 Million cell updates/sec

Title: US-10-010-667A-2_COPY_271_291

Perfect score: 93

Sequence: 1 WIDIKQFWYTPPTF 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	93	100.0	339 1	STEPA_HUMAN
2	52	55.9	458 1	AGLE_RHIME
3	45	48.4	865 1	ENV_HY1RH
4	43.5	46.8	1167 1	WCI_NEUCR
5	43	46.2	792 1	OSTA_XY1FT
6	42	45.2	346 1	OBSP_1CTPU
7	42	45.2	396 1	YOR1_YEAST
8	42	45.2	685 1	ERR2_YEAST
9	42	45.2	795 1	TLR1_MOUSE
10	42	45.2	847 1	ENV_HY1S1
11	41.5	44.6	224 1	YH85_YEAST
12	41	44.1	405 1	WCAD_ECOLI
13	41	44.1	430 1	PUCK_BACU
14	41	44.1	454 1	DEH4_SACBA
15	41	44.1	703 1	ARYB_MANSE
16	41	44.1	843 1	ENV_HY1Y2
17	41	44.1	854 1	ENV_STVCZ
18	41	44.1	856 1	ENV_HY1ZH
19	41	44.1	1509 1	MYSN_ACOCA
20	41	44.1	1628 1	NAGH_CLOPE
21	40	43.0	226 1	HA2P_RABIT
22	40	43.0	792 1	OSTA_XY1FA
23	40	43.0	847 1	ENV_HY1W2
24	40	43.0	855 1	ENV_HY1Z2
25	40	43.0	855 1	ENV_HY1Z6
26	40	43.0	856 1	ENV_HY1Z6
27	39.5	42.5	263 1	MP1_L0LPR
28	39.5	42.5	263 1	MP1_P0LPR
29	39.5	42.5	265 1	MEH1_H0LPA
30	39.5	42.5	269 1	MEH1_P0LPA
31	39.5	42.5	551 1	YABN_ECOLI
32	39.5	42.5	701 1	ACSA_HUMAN
33	39	41.9	210 1	SSH5_YEAST

34	39	41.9	223 1	CDR1_HUMAN	P51861 homo sapien
35	39	41.9	282 1	Y765_HAETN	Q57125 haemophilus
36	39	41.9	289 1	MTM3_METUA	Q58015 methanococc
37	39	41.9	361 1	SERC_BACHD	Q9K6M4 bacillus ha
38	39	41.9	459 1	NUAM_MACRO	P92684 macropus ro
39	39	41.9	847 1	OPGH_ECOLI6	O8F1S3 escherichia
40	39	41.9	847 1	OPGH_ECOLI	P33137 escherichia
41	39	41.9	847 1	OPGH_SALTY	Q82718 salmoneilla
42	39	41.9	847 1	OPGH_SALTY	Q82726 salmoneilla
43	39	41.9	853 1	ENV_HY1EL	P04581 human immun
44	39	41.9	867 1	OPGH_NITEU	O82888 nitrososoma
45	39	41.9	868 1	ENV_HY1C4	P05879 human immun

RESULT 1

STEPA_HUMAN	STANDARD;	ERR;	339 AA.
AC	Q9UHE8; 095034;		
DT	16-OCT-2001 (Rel. 40, Created)		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	15-MAR-2004 (Rel. 43, Last annotation update)		
DE	Six transmembrane epithelial antigen of prostate.		
GN	STEAP OR STEAP1.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=20056277; PubMed=10586738;		
RA	Hubert R.S., Vivanco I., Chen E., Rastegar S., Jeong K.,		
RA	Mitchell S.C., Madraswala R., Zhou J., Kuo J., Rattano A.B.,		
RA	Takobovits A., Saffran D.C., Afar D.B.H., antigen highly expressed in		
RT	"STEAP: a prostate-specific cell-surface antigen highly expressed in		
RT	human prostate tumors."		
RL	Proc. Natl. Acad. Sci. U.S.A. 96:14523-14528 (1999).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RA	Abu-Threideh J., Stoneking T., Langston Y., Maupin R.,		
RA	Submitted (JUN-1998) to the EMBL/Genbank/DBJ databases.		
RL	[3]		
RN	SEQUENCE FROM N.A.		
RP	TISSUE=Skin;		
RC	MEDLINE=2238257; PubMed=12477932;		
RX	Straussberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,		
RA	Altschul S.F., Zeeberg B., Buellow K.H., Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,		
RA	Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,		
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,		
RA	Brownstein M.J., Ussin T.B., Toshitsuki S., Carrinzi P., Prange C.,		
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,		
RA	Bosk S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Halyk S.W.,		
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,		
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,		
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,		
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,		
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,		
RA	Schneer A., Schein J.E., Jones S.J.M., Marra M.A.,		
RT	"Generation and initial analysis of more than 15,000 full-length		
RT	human and mouse cDNA sequences."		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).		
CC	- SUBCELLULAR LOCATION: Integral membrane protein (potential).		
CC	- TISSUE SPECIFICITY: Highly expressed in prostate tumors.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		

```
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF186249; AAF17479.1; -.
DR EMBL; AC005053; AAC7150.1; ALT_INIT.
DR EMBL; AC004969; AAD15620.2; -.
DR EMBL; BC011802; AAH11802.1; -.
DR Genew; HGNC:11378; STEAP.
DR MIM; 604415; -.
DR GO; GO:0005887; C:integral to plasma membrane; TAS.
DR GO; GO:0005911; C:intracellular junction; TAS.
DR GO; GO:0015267; F:channel/pore class transporter activity; TAS.
KM Transmembrane; Antigen.
FT TRANSMEM 71 91 POTENTIAL.
FT TRANSMEM 119 139 POTENTIAL.
FT TRANSMEM 164 184 POTENTIAL.
FT TRANSMEM 218 238 POTENTIAL.
FT TRANSMEM 258 278 POTENTIAL.
FT TRANSMEM 291 311 POTENTIAL.
SQ SEQUENCE 339 AA; 39851 MW; 55443A170C870387 CRC64;

Query Match
Best Local Similarity 100.0%; Score 93; DB 1; Length 339;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WIDIKQFWYTPPTF 15
DB 277 WIDIKQFWYTPPTF 291

RESULT 2
AGLE RHIME STANDARD; PRT; 458 AA.
AC Q923R5;
DT 30-MAY-2000 (Rel. 39, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Alpha-glucosides-binding periplasmic protein agle precursor.
GN AGLE OR R00695 OR SMC03061.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99328961; PubMed=10400573;
RA Willis L.B., Walker G.C.;
RT "A novel Sinorhizobium meliloti operon encodes an alpha-glucosidase
RT and a periplasmic-binding-protein-dependent transport system for
RT alpha-glucosides."
RT J. Bacteriol. 181:4176-4184(1999).
RL [2]
RN SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430;
RA Boistard P., Barloy-Hubler F., Goazy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godite T., Goffeau A., Kahn D., Kiss B., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galbert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021."
RT Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
RL - FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM
CC FOR ALPHA-GLUCOSIDES SUCH AS SUCROSE, MALTOSE AND TREHALOSE.
CC - SUBCELLULAR LOCATION: Periplasmic (Probable).
CC - SIMILARITY: Belongs to the bacterial extracellular solute-binding
CC protein family 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF045609; AAD12050.1; ALT_INIT.
DR EMBL; AL591784; CAC45267.1; -.
DR InterPro; IPR006059; SBP_dac.1.
DR InterPro; IPR006061; SBP_domi.
DR Pfam; PF01547; SBP_dac.1; 1.
DR PROSITE; PS01037; SBP_BACTERIAL_1; FALSE NEG.
KM Sugar transporter; Transport; Periplasmic; Signal; Complete proteome.
FT SIGNAL 1 27 POTENTIAL.
FT CHAIN 28 458 ALPHA-GLUCOSIDES-BINDING PERIPLASMIC
FT PROTEIN AGLE.
SQ SEQUENCE 458 AA; 49703 MW; E7A7F8157C2FC291 CRC64;

Query Match
Best Local Similarity 55.9%; Score 52; DB 1; Length 458;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 IDIKQFWYTPPTF 15
DB 173 IDIKSLWYVPENF 186

RESULT 3
ENV HVLRH STANDARD; PRT; 865 AA.
AC P04579;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contactus: Exterior membrane
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (RF/HAT isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11701;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86218077; PubMed=2423250;
RA Starchic B.R., Hahn B.H., Shaw G.M., McNeely P.D., Modrow S., Wolf H.,
RA Parks E.S., Parks W.P., Josephs S.F., Gallo R.C., Wong-Staal F.;
RT "Identification and characterization of conserved and variable
RT regions in the envelope gene of HTLV-III/LAV, the retrovirus of
RT AIDS."
RT Cell 45:637-648(1986).
RL -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M17451; AAA5057.1; -.
DR HIV; M17451; ENVSRF.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000773; GP120.
DR Pfam; PR00516; GP120; 1.
DR Pfam; PR00517; GP41; 1.
KM Aids; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
KW Signal.
FT SIGNAL 1 29
FT CHAIN 30 519 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 520 865 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 218 BY SIMILARITY.
FT DISULFID 125 209 BY SIMILARITY.
FT DISULFID 130 157 BY SIMILARITY.
```



```

FT DISULFID 231 260 BY SIMILARITY.
FT DISULFID 241 252 BY SIMILARITY.
FT DISULFID 309 343 BY SIMILARITY.
FT DISULFID 369 452 BY SIMILARITY.
FT DISULFID 396 425 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 129 129 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 139 139 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 193 193 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 194 194 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 203 203 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 210 210 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 247 247 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 254 254 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 275 275 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 302 302 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 308 308 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 314 314 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 344 344 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 351 351 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 367 367 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 403 403 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 407 407 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 413 413 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 418 418 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 455 455 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 471 471 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 620 620 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 625 625 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 634 634 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 646 646 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 825 825 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 865 AA; 97809 MM; 2882862314ADCC CRC64;

Query Match 48.4%; Score 45; DB 1; Length 865;
Best Local Similarity 60.0%; Pred. No. 13;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 WIDKQFWY 10
Db 681 WEDTQWLMY 690

RESULT 4
WCL_NEUCR STANDARD; PRT; 1167 AA.
ID WCL_NEUCR Q01371;
AC 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE White collar 1 protein (WCL).
GN WC-1.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OC NCBI_Taxid=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=74-OR23-1A / FGSC 987;
RX MEDLINE=96203083; PubMed=8612589;
RA Ballario P., Vitorioso P., Magrelli A., Talora C., Cabibbo A.,
RT "White collar-1, a central regulator of blue light responses in
RT Neurospora, is a zinc finger protein.";
RL EMBL J. 15:1650-1657(1996).
RN [2]
RP REVISIONS TO C-TERMINUS.

```

```

RA Ballario P.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MAY FUNCTION AS A TRANSCRIPTION FACTOR INVOLVED IN LIGHT
CC REGULATION. BINDS AND AFFECTS BLUE LIGHT REGULATION OF THE AL-3
CC GENE. WCL1 AND WCL2 PROTEINS INTERACT VIA HOMOLOGOUS PAS DOMAINS,
CC BIND TO PROMOTERS OF LIGHT REGULATED GENES SUCH AS FRQ, AND
CC ACTIVATE TRANSCRIPTION.
CC -!- SUBUNIT: HETERODIMER OF WCL1 AND WCL2 (POTENTIAL).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- INDUCTION: By blue light.
CC -!- DOMAIN: The glutamine-rich domain might function in activating
CC gene expression.
CC -!- SIMILARITY: Contains 1 GATA-type zinc finger.
CC -!- SIMILARITY: Contains 3 PAS (PBR-ARNT-SIM) dimerization domains.
CC -!- SIMILARITY: Contains 2 PAS-associated C-terminal (PAC) domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X94300; CAA63964.2; -.
DR HSRP; P17679; IGNP.
DR TRANSFAC; T02819; -.
DR InterPro; IPR001610; PAC.
DR InterPro; IPR000014; PAS domain.
DR InterPro; IPR00679; Znf_GATA.
DR Pfam; PF00320; GATA; 1.
DR Pfam; PF00785; PAC; 1.
DR Pfam; PF00989; PAS; 2.
DR SMART; SM00086; PAC; 2.
DR SMART; SM00091; PAS; 3.
DR SMART; SM00401; Znf_GATA; 1.
DR TIGRFAMs; TIGR00229; sensory_box; 3.
DR PROSITE; PS00344; GATA_ZN_FINGER_1; 1.
DR PROSITE; PS50114; GATA_ZN_FINGER_2; 1.
DR DR PROSITE; PS50112; PAS; 3.
KW Transcription regulation; Activator; DNA-binding; Zinc-finger;
KW Nuclear protein; Repeat.
FT DOMAIN 16 61 GIN-RICH.
FT FT 381 452 PAS 1.
FT FT 469 508 PAC 1.
FT FT 574 644 PAS 2.
FT FT 650 691 PAC 2.
FT FT 693 763 PAS 3.
FT FT 934 959 GATA-TYPE.
FT FT 21 57 POLY-GIN.
FT FT 329 333 POLY-PRO.
SQ SEQUENCE 1167 AA; 127454 MM; 6489D04DAB50EE38 CRC64;

Query Match 46.8%; Score 43.5; DB 1; Length 1167;
Best Local Similarity 66.7%; Pred. No. 32;
Matches 8; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 3 DIKQFWYTPPT 14
Db 529 DIQYIM-TPPT 539

RESULT 5
OSTA_XYLFT STANDARD; PRT; 792 AA.
ID OSTA_XYLFT Q87A19;
AC 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Organic solvent tolerance protein precursor.
GN IMP OR OSTA OR PD1836.
OS Xylella fastidiosa (strain Temecual / ATCC 700964).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;

```

```

OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=183190;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=2242131; PubMed=12533478;
RA Van Slyks M.A., de Oliveira M.C., Monteiro-Vitorello C.B., Moon D.H.,
RA Miyaki C.Y., Furlan L.R., Camargo L.E.A., da Silva A.C.R., Neeb D.H.,
RA Takita M.A., Lemos E.G.M., Machado M.A., Ferro M.I.T., da Silva F.R.,
RA Goldman M.B.S., Goldman G.H., Lemos M.V.F., El-Dorry H., Tsai S.M.,
RA Carrier H., Carraro D.M., de Oliveira R.C., Nunes L.R., Siqueira M.J.,
RA Coutinho L.L., Kimura E.T., Ferro E.S., Harakava R., Kuramae E.E.,
RA Maria C.L., Gigliotti E., Abreu I.L., Alves L.M.C., do Amaral A.M.,
RA Baio G.S., Blanco S.R., Brito M.S., Canavan F.S., Celestino A.V.,
RA da Cunha A.F., Fenille R.C., Ferro J.A., Formighieri E.F., Kishi L.T.,
RA Leoni S.G., Oliveira A.R., Rosa V.E. Jr., Sassaki F.T., Sena J.A.D.,
RA de Souza A.A., Trufil D., Tsukumo F., Yanai G.M., Zares L.G.,
RA Civerolo E.L., Simpson A.J.G., Almeida N.F. Jr., Setubal J.C.,
RA Kitajima J.P.;
RT "Comparative analyses of the complete genome sequences of Pierce's
RT disease and citrus variegated chlorosis strains of Xylella
RT fastidiosa."
RT J. Bacteriol. 185:1018-1026(2003).
CC -1- FUNCTION: Determines N-hexane tolerance. Involved in outer
CC membrane permeability. Essential for envelope biogenesis. Could be
CC part of a targeting/usher system for outer membrane components (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Outer membrane (By similarity).
CC -1- SIMILARITY: Belongs to the Imp/OscA family.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AE012560; AAO29668.1; -.
DR HAMAP; MF_01411; 1.
DR InterPro; IPR007543; OSCA_C.
DR Pfam; PF04453; OSCA_C; 1.
DR KX Outer membrane; Signal; Complete proteome.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 792 ORGANIC SOLVENT TOLERANCE PROTEIN.
SQ SEQUENCE 792 AA; 90768 MW; 70637D6FA7B62DAD CRC64;
Query Match 46.2%; Score 43; DB 1; Length 792;
Best Local Similarity 36.4%; Pred. No. 26;
Matches 8; Conservative 3; Mismatches 3; Indels 8; Gaps 1;
QY 2 IDIKOFV-----WYTPPTF 15
Db 440 IDVKPYVSLPFTGASWYTPTF 461

```

```

RT organ, defines a new gene family."
RL J. Neurosci. 17:8083-8092(1997).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: PARAPINEAL ORGAN.
CC -1- PTM: Some or all of the carboxyl-terminal Ser or Thr residues may
CC be phosphorylated (By similarity).
CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC Opstin subfamily.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AF028014; AAB84050.1; -.
DR HSSP; P02699; 1FDP.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR001760; Opstin.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS00237; G PROTEIN RECEPT_F1_1; 1.
DR PROSITE; PS00262; G PROTEIN RECEPT_F1_2; 1.
DR PROSITE; PS00238; OESIN; 1.
DR KX Photoreceptor; Retinal protein; Transmembrane; Glycoprotein;
DR Phosphorylation; Lipoprotein; Palmitate; G-protein coupled receptor.
FT DOMAIN 1 29
FT TRANSMEM 30 54
FT TRANSMEM 55 66
FT TRANSMEM 67 91
FT DOMAIN 92 106
FT TRANSMEM 107 126
FT TRANSMEM 127 145
FT TRANSMEM 146 169
FT DOMAIN 170 193
FT TRANSMEM 194 221
FT DOMAIN 222 244
FT TRANSMEM 245 268
FT TRANSMEM 269 276
FT TRANSMEM 277 301
FT DOMAIN 302 346
FT DISULFID 103 180
FT BINDING 288 288
FT CARBOHYD 8 8
FT CARBOHYD 191 191
FT LIPID 315 315
SQ SEQUENCE 346 AA; 38203 MW; A70871684F8FC7FD CRC64;
Query Match 45.2%; Score 42; DB 1; Length 346;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 7 FVWYTPPTP 15
Db 158 FVWYTPPTP 166

```

```

RESULT 6
ID OPSP ICTPU STANDARD; PRT; 346 AA.
AC 04226;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Parapinopsin.
OS Ictalurus punctatus (Channel catfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
OC Ictaluridae; Ictalurus.
OX NCBI_TaxID=7998;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=97477428; PubMed=9334384;
RA Blackshaw S., Snyder S.H.;
RT "Parapinopsin, a novel catfish opsin localized to the parapineal

```

```

RESULT 7
ID YJRL YEAST STANDARD; PRT; 396 AA.
AC P46992;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Hypothetical 43.0 kDa protein in CPS1-FP1 intergenic region.
GN YJL171C OR J0512.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN (1)

```

```
RP SEQUENCE FROM N.A.
RA Obermaier B., Piravandi E., Rinke M., Dondley H.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
CC
CC -1- SIMILARITY: TO YEAST YER162C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; 249446; CAA89466.1; -
DR PIR; S56954; S56954.
DR GerMonline; 141783; -.
DR SGD; S0003707; YOL171C.
DR Hypothetical protein.
SQ SEQUENCE 396 AA; 43014 MW; 279E858E7512670A CRC64;

Query Match      45.2%; Score 42; DB 1; Length 396;
Best Local Similarity 50.0%; Pred. No. 18;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 2 IDIKOFWWTPTPT 15
   :|||:|||
Db 83 LKXQFAFYTSBGF 96

RESULT 8
ERF2_YEAST
ID ERF2_YEAST STANDARD; PRT; 685 AA.
AC P05453; P05420;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Eukaryotic peptide chain release factor GTP-binding subunit (ERF2)
DE (Translational release factor 3) (ERF3) (ERF-3) (Omnipotent suppressor
DE protein 2) (G1 to S phase transition protein 1).
GN SUP35 OR SUP2 OR SUP12 OR GST1 OR SAL3 OR PMW2 OR YDR172W OR
GN YD9395.05.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_Taxid=4932;
KN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88329727; PubMed=3047009;
RA Kushnir V.V., Ter-Avanesyan M.D., Telckov M.V., Surguchov A.P.,
RT Smirnov V.N., Inge-Vecheromov S.G.;
RT "Nucleotide sequence of the SUP2 (SUP35) gene of Saccharomyces
RT cerevisiae.";
RL Gene 66:45-54 (1988).
RL [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87219095; PubMed=3556215;
RA Inge-Vecheromov S.G.;
RA Kushnir V.V., Ter-Avanesyan M.D., Surguchov A.P., Smirnov V.N.,
RT "Localization of possible functional domains in sup2 gene product of
RT the yeast Saccharomyces cerevisiae.";
RL FEBS Lett. 215:257-260 (1987).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=88172503; PubMed=3280807;
RA Wilson P.G., Culbertson M.R.;
RT "SUP12 suppressor protein of yeast. A fusion protein related to the
RT EF-1 family of elongation factors.";
RL J. Mol. Biol. 199:559-573 (1988).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=88296422; PubMed=2841115;
RA Kuchel Y., Shimatake H., Kikuchi A.;
RT "A yeast gene required for the G1-to-S transition encodes a protein
```

```
RT containing an A-kinase target site and GTPase domain.";
RL EMBL J. 7:1175-1182 (1986).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Murphy L., Harris D.E., Barrell B.G., Rajandream M.A.;
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
RN [6]
RP FUNCTION.
RX MEDLINE=96016209; PubMed=7556078;
RA Stanfield I., Jones K.M., Kushnir V.V., Dagkesamanskaya A.R.,
RA Poznyakovsky A.I., Pauskin S.V., Nierras C.R., Cox B.S.,
RA Ter-Avanesyan M.D., Tulce M.F.;
RT "The products of the SUP45 (erf1) and SUP35 genes interact to mediate
RT translation termination in Saccharomyces cerevisiae.";
RL EMBL J. 14:4365-4373 (1995).
CC -1- FUNCTION: INVOLVED IN TRANSLATION TERMINATION. STIMULATES THE
CC ACTIVITY OF ERF1. BINDS GUANINE NUCLEOTIDES.
CC -1- SUBUNIT: Heterodimer of two subunits, one of which binds GTP.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC ERF3 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M21129; AAA35133.1; -.
DR EMBL; X07163; CAA30155.1; -.
DR EMBL; Y00829; CAA68760.1; -.
DR EMBL; Z46727; CAA86677.1; -.
DR PIR; S00733; EFBYS2.
DR GerMonline; 140663; -.
DR SGD; S0002679; SUP35.
DR InterPro; IPR000795; EF_GTPbind.
DR InterPro; IPR004160; EFTU_Cterm.
DR InterPro; IPR004161; EFTU_D2.
DR InterPro; IPR009001; Elongat_init_C.
DR InterPro; IPR009000; Translat_factor.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR Pfam; PF03143; GTP_EFTU_D3; 1.
DR PRINTS; PR00315; ELONGATNPFCT.
DR PROSITE; PS00301; EFACOR_GTP; 1.
KW Protein biosynthesis; Repeat; Phosphorylation; GTP-binding.
FT DOMAIN 5 135
FT DOMAIN 139 249
FT NP_BIND 267 274 GTP (BY SIMILARITY).
FT NP_BIND 344 348 GTP (BY SIMILARITY).
FT NP_BIND 406 409 GTP (BY SIMILARITY).
FT ACT_SITE 273 273 INTERACT WITH GTP/GDP (BY SIMILARITY).
FT ACT_SITE 407 407 INTERACT WITH GTP/GDP (BY SIMILARITY).
FT MOD_RES 341 341 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 341 341 S -> C (IN REF. 4).
SQ SEQUENCE 685 AA; 76551 MW; 43912A6D7DFA153 CRC64;

Query Match      45.2%; Score 42; DB 1; Length 685;
Best Local Similarity 53.8%; Pred. No. 32;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 2 IDIKOFWWTPTPT 14
   :|||:|||
Db 459 VDPKCEPWTGPT 471

RESULT 9
TIR1_MOUSE
ID TIR1_MOUSE STANDARD; PRT; 795 AA.
AC Q9ERP1; Q9EPW5;
```

DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Toll-like receptor 1 precursor (Toll/interleukin-1 receptor-like)
DE (TL1).
GN TL1.
OS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId:10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Macrophage;
RX MEDLINE=20558581; PubMed=11095740;
RA Ozinsky A., Underhill D.M., Fontenot J.D., Hajjar A.M., Smith K.D.,
RA Wilson C.B., Schroeder L., Aderem A.;
RT "The repertoire for pattern recognition of pathogens by the innate
RT immune system is defined by cooperation between Toll-like
RT receptors.";
RT Proc. Natl. Acad. Sci. U.S.A. 97:13766-13771 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Macrophage;
RX MEDLINE=20571875; PubMed=11123271;
RA Hajjar A.M., O'Mahony D.S., Ozinsky A., Underhill D.M., Aderem A.,
RA Klebanoff S.J., Wilson C.B.;
RT "Functional interactions between Toll-like receptor (TLR) 2 and TLR1
RT or TLR6 in response to phenol-soluble modulins.";
RN J. Immunol. 166:15-19 (2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Spleen;
RA Thomson D.P., Campbell C.C., Liew F.Y., Xu D.;
RT "Cloning of Mus musculus Toll-like receptor 1.";
RT Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Participates in the innate immune response to microbial
CC agents. Cooperates with TLR2 and modulates the response to
CC microbial constituents. Acts via MyD88 and TRAF6, leading to NF-
CC kappa-B activation, cytokine secretion and the inflammatory
CC response (By similarity).
CC -1- SUBUNIT: Binds TLR2 via their respective extracellular domains.
CC Binds MyD88 via their respective TIR domains (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Plasma membrane and
CC plasmomes.
CC -1- SIMILARITY: Belongs to the Toll-like receptor family.
CC -1- SIMILARITY: Contains 1 TIR domain.
CC -1- SIMILARITY: Contains 9 leucine-rich (LRR) repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial/
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC -----
DR EMBL; AY009154; AAC37302.1; -
DR EMBL; AF316985; AAG35062.1; -
DR HSSP; O60603; 1PYW.
DR MGD; MGI:1341295; T1rl.
DR GO; GO:0016020; C:membrane; NAS.
DR GO; GO:0045335; C:phagocytic vesicle; NAS.
DR GO; GO:0003793; F:defense/immunity protein activity; NAS.
DR GO; GO:0004888; F:transmembrane receptor activity; NAS.
DR GO; GO:00042497; F:triacylated lipoprotein binding; NAS.
DR GO; GO:0007250; P:activation of NF-kappa-B-inducing kinase; NAS.
DR GO; GO:0043216; P:macrophage activation; NAS.
DR GO; GO:0042495; P:perception of triacylated bacterial lipopo.
DR GO; GO:0045410; P:positive regulation of interleukin-6 biosyn.
DR GO; GO:0042535; P:positive regulation of tumor necrosis facto.
DR InterPro; IPR004075; IIL_receptortl.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.

DR	InterPro; IPR000157; TIR.
DR	Pfam; PF00560; LRR; 8.
DR	Pfam; PF01463; LRCT; 1.
DR	Pfam; PF01582; TIR; 1.
DR	PRINTS; PR01537; INTRLNARF.
DR	PRINTS; PR00019; LEURICHRPT.
DR	SMART; SMO0082; LRCT; 1.
DR	SMART; SMO0255; TIR; 1.
DR	PROSITE; PS50104; TIR; 1.
KW	Receptor; Immune response; Inflammatory response; Signal; Transmembrane; Repeat; Leucine-rich repeat; Glycoprotein. POTENTIAL.
FT	SIGNAL 1 25
FT	CHAIN 26 795
FT	DOMAIN 26 582
FT	TRANSMEM 583 603
FT	DOMAIN 604 795
FT	REPEAT 46 70
FT	REPEAT 71 94
FT	REPEAT 96 116
FT	REPEAT 117 140
FT	REPEAT 374 397
FT	REPEAT 400 427
FT	REPEAT 447 470
FT	REPEAT 471 494
FT	REPEAT 496 515
FT	DOMAIN 638 782
FT	CARBOHYD 38 38
FT	CARBOHYD 59 59
FT	CARBOHYD 88 88
FT	CARBOHYD 140 140
FT	CARBOHYD 166 166
FT	CARBOHYD 251 251
FT	CARBOHYD 296 296
FT	CARBOHYD 333 333
FT	CARBOHYD 432 432
FT	CONFLICT 88 88
SO	SEQUENCE 795 AA; 90672 MW; 855356429872D232 CRC64;
Query Match 45.2%; Score 42; DB 1; Length 795;	
Best Local Similarity 58.3%; Pred. No. 38;	
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;	
Dy	1 WIDIKQFWYTP 12 :: ::
Dd	264 FINIIQIVWHP 275
RESULT 10	
ENV_HV1S1	STANDARD; PRT; 847 AA.
AC	P19550;
DT	01-FEB-1991 (Rel. 17, Created)
DT	01-FEB-1991 (Rel. 17, Last sequence update)
DT	10-OCT-2003 (Rel. 42, last annotation update)
DE	Envelope polypeptide GP160 precursor [contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
DN	ENV.
OS	Human immunodeficiency virus type 1 (SF162 isolate) (HIV-1).
OC	Vituses; Retroid vituses; Retroviridae; Lentivirns.
OX	NCBI_Taxid=11691;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=90347835; PubMed=2384920;
RA	Cheng-Mayer C., Quiroga M., Tung J.W., Dina D., Levy J.;
RT	"Viral determinants of human immunodeficiency virus type 1 T-cell or macrophage tropism, cytopathogenicity, and CD4 antigen modulation.";
RL	J. Virol. 64:4390-4398(1990).
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; M65024; AAA5072.1; -

DR PDB; 10BE; 15-MAY-97.

DR HIV; M38428; ENV5SF162.

DR InterPro; IPR000328; Env GP41.

DR InterPro; IPR000777; GP120.

DR Pfam; PF00516; GP120; 1.

DR Pfam; PF00517; GP41; 1.

KW Aids; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal; 3D-structure.

FT SIGNAL 1 29

FT CHAIN 30 502

FT CHAIN 503 847

FT DISULFID 53 73

FT DISULFID 118 203

FT DISULFID 125 194

FT DISULFID 130 155

FT DISULFID 216 245

FT DISULFID 226 237

FT DISULFID 294 328

FT DISULFID 374 435

FT DISULFID 381 408

FT CARBOHYD 87 87

FT CARBOHYD 135 135

FT CARBOHYD 154 154

FT CARBOHYD 186 186

FT CARBOHYD 195 195

FT CARBOHYD 232 232

FT CARBOHYD 239 239

FT CARBOHYD 260 260

FT CARBOHYD 274 274

FT CARBOHYD 293 293

FT CARBOHYD 299 299

FT CARBOHYD 329 329

FT CARBOHYD 336 336

FT CARBOHYD 352 352

FT CARBOHYD 382 382

FT CARBOHYD 388 388

FT CARBOHYD 392 392

FT CARBOHYD 398 398

FT CARBOHYD 401 401

FT CARBOHYD 438 438

FT CARBOHYD 454 454

FT CARBOHYD 602 602

FT CARBOHYD 607 607

FT CARBOHYD 616 616

FT CARBOHYD 628 628

SEQ SEQUENCE 847 AA; 96135 MW; 0A901317PD7FF2AB CRC64;

Query Match 45.2%; Score 42; DB 1; Length 847;

Best Local Similarity 50.0%; Pred. No. 40;

Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 MIDIKOFVWY 10

DB 663 WFDISKLMY 672

RESULT 11

YN89 YEAST

ID YN89 YEAST STANDARD; PRT; 224 AA.

AC P53721;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 01-OCT-1996 (Rel. 34, Last annotation update)

DE Hypothetical 25.3 kDa protein in TIM3-ARE2 intergenic region.

GN YN018W OR N185.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

NCBI_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A.

RA Pohl T.M.;

RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC

DR EMBL; 271633; CAA96297.1; -

DR PIR; S63349; S63349.

DR GeronOnline; 143363; -

DR SGD; S0005301; YNR018W.

DR InterPro; IPR007667; HIG_1_N.

DR Pfam; PF04588; HIG_1_N; 1.

KW Hypothetical protein; Transmembrane.

FT TRANSMEM 19 39

FT TRANSMEM 50 70

FT TRANSMEM 152 172

FT DOMAIN 196 199

SEQ SEQUENCE 224 AA; 25344 MW; FA2C528A008CE7C CRC64;

Query Match 44.6%; Score 41.5; DB 1; Length 224;

Best Local Similarity 57.1%; Pred. No. 12;

Matches 8; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

OY 1 MIDIKOFVWYPT 14

DB 52 W-SIKTALMTPPT 64

RESULT 12

WCAD ECOLI

ID WCAD ECOLI STANDARD; PRT; 405 AA.

AC P71238; P76385;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Putative colanic acid polymerase.

GN WCAD OR B2056.

OS Escherichia coli.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.

NCBI_TaxID=562;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=K12;

RX MEDLINE=96326333; PubMed=8759852;

RA Stevenson G., Andrianopoulos K., Hobbs M., Reeves P.R.;

RT "Organization of the Escherichia coli K-12 gene cluster responsible for production of the extracellular polysaccharide colanic acid.";

RL J. Bacteriol. 178:4885-4893 (1996).

RN [2]

RP REVISIONS.

RC STRAIN=K12;

RA Reeves P.R.;

RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=K12 / MG1655;

RX MEDLINE=97426617; PubMed=9279503;

RA Bistner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y.;

RA "The complete genome sequence of Escherichia coli K-12.";

RT Science 277:1453-1474 (1997).

RL -1- PATHWAY: Slime polysaccharide colanic acid biosynthesis.

```
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U38473; AAC77839.1; -.
CC EMBL: AB000295; AAC75117.1; -.
CC PIR: G64971; G64971.
CC Ecocyc: EG13572; wcad.
CC Lipopolysaccharide biosynthesis; Complete proteome.
CC CONFLICT 108 45409 MW; A3D9D91255686043 CRC64;
SQ SEQUENCE 405 AA; 45409 MW; A3D9D91255686043 CRC64;

Query Match
Best Local Similarity 57.1%; Score 41; DB 1; Length 405;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 WDIKQFWYTPPT 14
Db 188 WLSIKQFGIKTPT 201

RESULT 13
PUCB BACSU STANDARD; PRT; 430 AA.
AC 032140;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Uric acid permease pucb.
GN PUCB OR BS032440.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
ON NCBI_TaxID=1423;
RX MEDLINE=9804403; PubMed=9384377;
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borris R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denzot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Eutanz K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallier A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Haeck J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koeltter P., Konigstein G., Krog S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetlelle D., Rapoport G., Rey M., Reynolds S.,
RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serron P., Shin B.S., Soldo B.,
RA Sorokin A., Taccoti E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandemol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegeger T.,
RA Winers P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis."
RT Nature 390:249-256(1997).
```

```
RN [2]
RP FUNCTION:
RC STRAIN=168;
RX MEDLINE=21242727; PubMed=11344136;
RA Schults A.C., Mygaard P., Saxild H.H.;
RT "Functional analysis of 14 genes that constitute the purine catabolic
RT pathway in Bacillus subtilis and evidence for a novel region
RT controlled by the PucR transcription activator."
RL J. Bacteriol. 183:3293-3302(2001).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (probable).
CC -1- INDUCTION: Expression is very low in excess nitrogen (glutamate
CC (glutamate) and is induced during limiting nitrogen conditions
CC (glutamate). Expression is further induced when allantoin or uric
CC acid are added during limiting nitrogen conditions.
CC -1- SIMILARITY: BELONGS TO THE XANTHINE/URACIL PERMEASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z99120; CAB15234.1; -.
CC PIR: F70016; F70016.
CC Subtilist; BG13985; pucb.
CC InterPro: IPR006042; Xan_ur permease.
CC InterPro: IPR006043; Xan/urac/vitc.
CC Pfam: PF00860; xan_ur permease; 1.
CC TIGRFAMs: TIGR00801; ncs2; 1.
CC PROSITE: PS01116; XANTH URACIL PERMEASE; 1.
CC Transmembrane; Transport; Complete proteome.
GN TRANSMEM 18
FT TRANSMEM 18
FT TRANSMEM 43
FT TRANSMEM 67
FT TRANSMEM 97
FT TRANSMEM 122
FT TRANSMEM 153
FT TRANSMEM 185
FT TRANSMEM 209
FT TRANSMEM 233
FT TRANSMEM 274
FT TRANSMEM 310
FT TRANSMEM 333
FT TRANSMEM 369
FT TRANSMEM 398
SQ SEQUENCE 430 AA; 44861 MW; 9F3D18D100A0572F CRC64;

Query Match
Best Local Similarity 33.3%; Score 41; DB 1; Length 430;
Matches 5; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 WDIKQFWYTPPT 15
Db 215 WLHVSLEFPGPTF 229

RESULT 14
DH4 SACEA STANDARD; PRT; 454 AA.
ID DH4 SACEA
AC 08TF6;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE NADP-specific glutamate dehydrogenase I (EC 1.4.1.4) (NADP-GDH 1).
GN GDH1.
OS Saccharomyces bayanus (Yeast) (Saccharomyces uvarum).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
ON NCBI_TaxID=4931;
RX [1]
```


This Page Blank (uspto)

Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 14:01:53 ; Search time 7.5 Seconds

(without alignments)
192.383 Million cell updates/sec

Title: US-10-010-667A-2_COPY_271_291

Perfect score: 93

Sequence: 1 WIDIKQFWYTPPTF 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	48	51.6	453	2	A97431
2	48	51.6	453	2	AB2649
3	45	48.4	31	2	A23341
4	45	48.4	358	2	S21998
5	44.5	47.8	553	2	AC0065
6	44.5	47.8	662	2	E69274
7	44	47.3	331	2	AE3544
8	44	47.3	361	2	AH1469
9	44	47.3	425	2	T27241
10	44	47.3	568	2	AB9958
11	43.5	46.8	1154	2	S69206
12	43	46.2	402	2	T14669
13	43	46.2	418	2	T14946
14	43	46.2	863	2	AS5034
15	43	46.2	877	2	S49197
16	42	45.2	133	2	G81710
17	42	45.2	321	2	E81411
18	42	45.2	396	2	SS6954
19	42	45.2	462	2	B86262
20	42	45.2	639	2	E73336
21	42	45.2	685	1	EBFYS2
22	41.5	44.6	126	2	G69990
23	41.5	44.6	224	2	S63349
24	41	44.1	254	2	D69140
25	41	44.1	357	2	S21992
26	41	44.1	359	2	T44816
27	41	44.1	368	2	P84300
28	41	44.1	405	2	E90986
29	41	44.1	405	2	H85831

ALIGNMENTS

30	41	44.1	405	2	G64971	Purative colanic a
31	41	44.1	430	2	F70016	putine permease ho
32	41	44.1	703	2	B34434	aryphorin beta ch
33	41	44.1	843	1	H44001	env polyprotein pr
34	41	44.1	847	2	T09448	env polyprotein pr
35	41	44.1	847	2	S13289	env polyprotein pr
36	41	44.1	852	2	T12016	env polyprotein pr
37	41	44.1	854	1	VCLJST	env polyprotein pr
38	41	44.1	856	1	A44963	env polyprotein pr
39	41	44.1	929	2	G81036	env polyprotein pr
40	41	44.1	929	2	B81982	env polyprotein pr
41	41	44.1	1042	2	S43904	env polyprotein pr
42	41	44.1	1509	1	A27224	env polyprotein pr
43	40.5	43.5	313	2	A12642	env polyprotein pr
44	40.5	43.5	313	2	B97425	env polyprotein pr
45	40.5	43.5	1207	2	T00378	env polyprotein pr

RESULT 1
A97431
alpha-glucosides-binding periplasmic protein aglE precursor [imported] - Agrobacterium
C/Species: Agrobacterium tumefaciens
C/Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
C/Accession: A97431
R/Goodner, B.; Hinkle, G.; Gattling, S.; Miller, N.; Blanchard, M.; Ourliot, B.; Goldman
A.; Liu, F.; Wollam, G.; Gattling, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, B
Science 294, 2323-2328, 2001
A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tu
A/Reference number: A97359; WUID:21608551; PMID:11743194
A/Accession: A97431
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-453 <KID>
A/Cross-references: GB:AB007869; PIDN:AAK6402.1; PID:G15155534; GSPDB:GN00169
C/Genetics:
A/Gene: AGR_C 1045
A/Map position: circular chromosome

Query Match 51.6%; Score 48; DB 2; Length 453;
Best Local Similarity 53.8%; Pred. No. 4.1;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 DIKQFWYTPPTF 15
DB 169 DKSLVWVPENF 181

RESULT 2
AB2649
hypothetical protein aglE [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C/Species: Agrobacterium tumefaciens
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C/Accession: AB2649
R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo,
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kueyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A/Authors: Yoo, H.; Tso, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A/Reference number: AB2577; WUID:21608550; PMID:11743193
A/Accession: AB2649
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-453 <KID>
A/Cross-references: GB:AE008688; PIDN:AAL1608.1; PID:G17738945; GSPDB:GN00186
A/Experimental source: strain C58 (Dupont)
C/Genetics:
A/Gene: aglE
A/Map position: circular chromosome

Query Match	51.6%	Score 48	DB 2	Length 453
Best Local Similarity	53.8%	Pred. No. 4.1		
Matches	7	Conservative	1	Mismatches 5; Indels 0; Gaps 0;
QY	3	DIKQFWWTPTPTF	15	
	:			
Db	169	DVKSLWIYVPEPNF	181	

```

RESULT 3
A23341
Allergen R7 - perennial ryegrass (tentative sequence) (fragment)
C;Species: Lolium perenne (perennial ryegrass)
C;Date: 05-Jun1997 #sequence_revision 05-Jun-1987 #text_change 09-Jun-2000
C;Accession: A23341
R;Cottam, G.P.; Moran, D.M.; Standing, R.
Biochem. J. 234, 305-310, 1986
A;Title: Physicochemical and immunochemical characterization of allergenic proteins from
A;Reference number: A23341; MUID:86242068; PMID:3718469
A;Accession: A23341
A;Molecule type: protein
A;Residues: 1-51 <COT>
C;Superfamily: expansin

```

Query Match	48.4%	Score 45	DB 2	Length 31
Best Local Similarity	50.0%	Pred. No. 0.65		
Matches	7	Conservative	1	Mismatches 6
				Indels 0
				Gaps 0
QY	1	WIDIKQVWYRPT	14	
		: : : : :		
Db	18	WIDAKSTWVGKPT	31	

RESULT 4
S21998
envelope protein gp120/gp41 - human immunodeficiency virus type 1
C/Species: human immunodeficiency virus type 1, HIV-1
C/Variety: isolate 28
C/Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 01-Dec-2000
C/Accession: S21998; S70425
R/Steuier, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.
submitted to the EMBL Data Library, July 1991
A/Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determined by polymerase chain reaction
A/Reference number: S21990
A/Accession: S21998
A/Molecule type: DNA
A/Residues: 1-358 <STE1>
A/Cross-references: EMBL:X61359; NID:G60182; PIDN:CAA43630.1; PID:G60183
R/Steuier, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A>Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid
A/Reference number: S70417; MUID:92144209; PMID:11736940
A/Accession: S70425
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-222, 'X', 224-358 <STE2>
A/Cross-references: EMBL:X61359; NID:G60182; PIDN:CAA43630.1; PID:G60183
A/Superfamily: type B retrovirus env polyprotein

		Score 45; DB 2;	Length 358;
Query March		Pred. No. 9.9;	
Best Local Similarity		60.0%	
Matches	6; Conservative	2; Mismatches	2; Indels
			Gaps 0.
QY	I WIDIKQPVWY 10		
Db	174 WFDTITQMLWY 183		

RESULT 5
AC0065
conserved hypothetical protein YP00524 [imported] - *Yersinia pestis* (strain CO92)
C:Species: *Yersinia pestis*

C:/date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C:/Accession: ACO0065
R:/Parkhill, J., Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
deno-Tarrag, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
ll, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrrell,
Nature 413, 523-527, 2001
A:/Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
A:/Reference number: AB0001; MUID:21470413; PMID:11586360
A:/Accession: ACO065
A:/Status: Preliminary
A:/Molecule type: DNA
A:/Residues: 1-553 <KUR>
A:/Cross-references: GB:AL590842; PIDW:CAC89382.1; PID:gl5978619; GSPDB:GN00175
A:/Genetics:
A:/Gene: YPO0524

Query	Match	Similarity	Score	DB	Length	553;
Best Local	Similarity	53.8%				
Matches	7; Conservative	1;	Mismatches	4;	Indels	1;
Qy	1	WIDIKQEVWYRPP	13			
Db	540	WEDFKS-AMFTDP	551			

RESULT 6
E69274
acetyl-CoA synthetase (acs-1) homolog - Archaeoglobus fulgidus
C/Species: Archaeoglobus fulgidus
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 21-Jan-2000
C/Accession: E69274
R/Klank, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
..; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glock, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A/Authors: Utterback, T.; Cotton, M.D.; Springs, T.; Artich, P.; Kaime, B.P.; Sykes, S.
Smith, H.O.; Woesse, C.R.; Venter, J.C.
A/Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A/Reference number: A69250; PMID:98049343; PMID:9389475
A/Accession: E69274
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-662 <KIR>
A/Cross-references: GB:AE001092; GB:AE000782; NID:g2689415; PIDN:AA891033.1; PID:g2650444
C/Superfamily: acetate-CoA ligase; acetate-CoA ligase homology
/145-621/Domain: acetate-CoA ligase homology <ACL>

	Query Match	47.8%	Score 44.5;	DB 2,	Length 662;
	Best Local Similarity	41.2%	Pred. No. 24;		
Matches	7; Conservative	6; Mismatches	1; Indels	3; Gaps	1
QY	2 IDIK--FVNYTPPTF 15				
	: : : : : :				
Db	307 MDFKGGDFLWYSPSW 323				

RESULT 7
 AF3544
 Hypothetical membrane spanning protein BMEI10279 [imported] - *Brucella melitensis* (strain
 C/Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
 C/Accession: AF3544
 R/Accession: AF3544
 R/Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
 R/Author: Kapral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
 M.; Mazur, M.; Goldsman, E.; Salkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letesche,
 Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
 A/Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*
 A/Reference number: AD5252; PMID:11756688
 A/Accession: AF3544
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-331 <KUR>
 A/Cross-references: GB:AE008918; PIDN:AAL53521.1; PID:G17984427; GSPDB:GM00191
 A/Experimental source: Strain 16M

C:Genetics:
A:Gene: BMEI10279
A:Map position: 11

Query Match 47.3%; Score 44; DB 2; Length 331;
Best Local Similarity 38.5%; Pred. No. 13;
Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 3 DIKQFWYTPPT 15
DB 117 NVKDFWHTPDP 129

RESULT 8

A:Accession: AH1469

Internalin protein homolog lin0295 [imported] - *Listeria innocua* (strain C1p11262)C:Species: *Listeria innocua*

C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001

C:Accession: AH1469

R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bioecker
; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Fsihl, H.
D.; Jones, L.M.; Kars, U.
Science 294, 849-852, 2001

A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapk, G.; Madueno, E.; Maitournam, A.; Ma

ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,

A:Title: Comparative genomics of *Listeria* species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AH1469

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-361 <GLA>

A:Cross-references: GB:AL592022; PIDN:GAC95528.1; PID:G16412724; GSPDB:GN00178

A:Experimental source: strain C1p11262

C:Genetics:

A:Gene: lin0295

Query Match 47.3%; Score 44; DB 2; Length 361;
Best Local Similarity 42.9%; Pred. No. 15;

Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 IDIKQFWYTPPT 15
DB 310 VTIHDMWYTPGT 323

RESULT 9

T27241

hypothetical protein Y57G11C.31 - *Caenorhabditis elegans*C:Species: *Caenorhabditis elegans*

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T27241

R:McMurray, A.
Submitted to the EMBL Data Library, September 1997

A:Reference number: Z20330

A:Accession: T27241

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-425 <WII>

A:Cross-references: EMBL:Z99281; PIDN:CA616530.1; GSPDB:GN00022; CESP:Y57G11C.31

A:Experimental source: clone Y57G11C

C:Genetics:

A:Gene: CESP:Y57G11C.31

A:Map position: 4

A:introns: 160/3; 187/2; 309/2; 361/3

Query Match 47.3%; Score 44; DB 2; Length 425;
Best Local Similarity 46.7%; Pred. No. 18;

Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 WIDIKQFWYTPPT 15
DB 146 WPEIQFPMPTPSLY 160

RESULT 10

A69958

acetyl-CoA synthetase [imported] - *Staphylococcus aureus* (strain N315)C:Species: *Staphylococcus aureus*

C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 17-May-2002

C:Accession: A69958

R:Kuroda, M.; Ohta, T.; Uchiyama, T.; Baba, T.; Yuzawa, H.; Kobayashi, T.; Cui, L.; Ogu
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.,
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.

A:Reference number: A69758; MUID:21311952; PMID:11418146

A:Accession: A69958

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-568 <KUR>

A:Cross-references: GB:BA000018; PID:G13701528; PIDN:BA42822.1; GSPDB:GN00149

A:Experimental source: strain N315

C:Genetics:

A:Gene: acsA

A:Superfamily: acetate-CoA ligase; acetate-CoA ligase homology

Query Match 47.3%; Score 44; DB 2; Length 568;
Best Local Similarity 58.3%; Pred. No. 24;

Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 DIKQFWYTPPT 14
DB 292 DFKVTIWTAPT 303

RESULT 11

S69206

regulator protein white collar 1 - *Neurospora crassa*C:Species: *Neurospora crassa*

C:Date: 21-Apr-1997 #sequence_revision 09-May-1997 #text_change 11-Jan-2002

C:Accession: S69206

R:Ballarino, P.; Vittorioso, P.; Magrelli, A.; Talora, C.; Cabibbo, A.; Macino, G.
EMBO J. 15, 1650-1657, 1996

A:Title: White collar-1, a central regulator of blue light responses in *Neurospora*, is

A:Reference number: S69206; MUID:96203083; PMID:8612589

A:Accession: S69206

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1154 <BAL>

A:Cross-references: EMBL:X94300; NID:G1279576; PID:G1480115

C:Genetics:

A:introns: 967/3

C:Superfamily: GATA-type zinc finger homology

C:Keywords: zinc finger

F:932-991/Domain: GATA-type zinc finger homology <GZF>

Query Match 46.8%; Score 43.5; DB 2; Length 1154;
Best Local Similarity 66.7%; Pred. No. 65;

Matches 8; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 3 DIKQFWYTPPT 14
DB 529 DIGQYIW-TPPT 539

RESULT 12

T14669

P-100 protein - *Yersinia pestis* plasmid pMT1C:Species: *Yersinia pestis*

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T14669

R:Hu, P.; Elliott, J.; McCreedy, P.; Skowronski, E.; Garnea, J.; Kobayashi, A.; Carrano
submitted to the EMBL Data Library, March 1998

A:Description: Structural organization of virulence determinants in three *Yersinia pestis*

A:Reference number: Z18168

A:Accession: T14669

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-402 <HUP>
A:Cross-References: EMBL:AF053947; NID:G2996286; PID:G2996304; PIDN:AA013184.1
C:Genetics:
A:Genome: plasmid PMT1

Query Match 46.2%; Score 43; DB 2; Length 402;
Best Local Similarity 63.6%; Pred. No. 24;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 KQFWYTPPTF 15
|||
Db 38 KORVWVAPTY 48

RESULT 13

hypothetical protein Y1030 - Yersinia pestis plasmid PMT1
C:Species: Yersinia pestis
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #ext_change 22-Oct-1999
C:Accession: T14946
R:Indler, L.E.; Plano, G.V.; Burland, V.; Mayhew, G.F.; Blattner, F.R.
Infect. Immun. 66, 5731-5742, 1998
A:Title: Complete DNA sequence and detailed analysis of the Yersinia pestis KIMS plasmid
A:Reference number: Z18268; MUID:99043898; PMID:9826348
A:Accession: T14946
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-418 <LIN>
A:Cross-References: EMBL:AF074611; NID:G3883003; PID:G3883031; PIDN:AA02691.1
C:Genetics:
A:Gene: Y1030
A:Genome: plasmid PMT1

Query Match 46.2%; Score 43; DB 2; Length 418;
Best Local Similarity 63.6%; Pred. No. 25;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 KQFWYTPPTF 15
|||
Db 54 KORVWVAPTY 64

RESULT 14

gag polyprotein - human immunodeficiency virus type 1 (strain Ant70)
A:Accession: A53034
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 03-May-1994 #sequence_revision 03-May-1994 #ext_change 07-May-1999
R:Vanden Haesevelde, M.; Decourt, J.L.; De Leys, R.J.; Vanderborcht, B.; van der Groen, U. Virol. 68, 1586-1596, 1994
A:Title: Genomic cloning and complete sequence analysis of a highly divergent African hu
A:Reference number: A53034; MUID:94149849; PMID:8107220
A:Accession: A53034
A:Status: preliminary
A:Molecule type: genomic RNA
A:Residues: 1-863 <VAN>
A:Cross-References: GB:L02587
C:Superfamily: type E retrovirus env polyprotein
C:Keywords: polyprotein

Query Match 46.2%; Score 43; DB 2; Length 863;
Best Local Similarity 50.0%; Pred. No. 57;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 WIDIKQVWY 10
|:|:|:|:|
Db 672 WIDITKMLWY 681

RESULT 15

S49197

envelope protein precursor - human immunodeficiency virus type 1 (fragment)
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 19-Mar-1997 #sequence_revision 24-Jul-1997 #ext_change 26-Aug-1999
C:Accession: S49197
R:Charneau, P.; Borman, A.M.; Quillent, C.; Guetard, D.; Charet, S.; Cohen, J.; Remy, submitted to the EMBL Data Library, July 1994
A:Description: Isolation and envelope sequence of a highly divergent HIV-1 isolate; defi
A:Reference number: S49197
A:Accession: S49197

Query Match 46.2%; Score 43; DB 2; Length 877;
Best Local Similarity 50.0%; Pred. No. 58;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 WIDIKQVWY 10
|||
Db 686 WIDITKMLWY 695

Search completed: July 12, 2004, 14:11:20
Job time : 8.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 14:10:33 ; Search time 26.5116 Seconds

(Without alignments)
176,480 Million cell updates/sec

Title: US-10-010-667a-2_COPY_277_291

Sequence: 1 WIDIKQFWYTPPTF 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1279676 seqs, 311918243 residues

Total number of hits satisfying chosen parameters: 1279676

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	93	100.0	15	14	US-10-011-095-21
2	93	100.0	15	14	US-10-010-667A-21
3	93	100.0	15	14	US-10-165-044-39
4	93	100.0	267	9	US-09-747-835A-50
5	93	100.0	267	9	US-09-747-835A-51
6	93	100.0	267	12	US-10-312-312-50
7	93	100.0	267	12	US-10-312-312-51
8	93	100.0	339	9	US-09-759-143-879
9	93	100.0	339	9	US-09-780-669-879
10	93	100.0	339	9	US-09-822-827-879
11	93	100.0	339	9	US-09-802-520-11
12	93	100.0	339	9	US-09-895-793-879
13	93	100.0	339	9	US-09-895-814-879
14	93	100.0	339	12	US-10-205-267-13
15	93	100.0	339	12	US-10-408-009-2

16	93	100.0	339	13	US-10-012-896-879	Sequence 879, App
17	93	100.0	339	14	US-10-011-095-21	Sequence 2, Appl1
18	93	100.0	339	14	US-10-010-667A-2	Sequence 2, Appl1
19	93	100.0	339	14	US-10-205-823-397	Sequence 397, App
20	93	100.0	339	14	US-10-144-678A-879	Sequence 879, App
21	93	100.0	339	14	US-10-294-025-879	Sequence 37, Appl
22	93	100.0	339	15	US-10-239-607-37	Sequence 37, Appl
23	93	100.0	339	15	US-10-295-027-714	Sequence 714, App
24	93	100.0	339	15	US-10-295-027-1347	Sequence 1347, Ap
25	93	100.0	368	12	US-10-425-114-72779	Sequence 72779, A
26	93	100.0	375	14	US-10-165-044-2	Sequence 2, Appl1
27	45	48.4	268	9	US-09-854-816-39	Sequence 39, Appl
28	45	48.4	269	9	US-09-854-816-33	Sequence 33, Appl
29	45	48.4	483	12	US-10-282-122A-57398	Sequence 57398, A
30	45	48.4	854	14	US-10-369-294-17	Sequence 17, Appl
31	45	48.4	871	15	US-10-346-000A-9	Sequence 9, Appl1
32	44.5	47.8	553	12	US-10-282-122A-77905	Sequence 77905, A
33	44.5	47.8	662	15	US-10-369-493-921	Sequence 921, App
34	44	47.3	70	16	US-10-437-963-139012	Sequence 139012, A
35	44	47.3	91	12	US-10-437-963-115402	Sequence 115402, A
36	44	47.3	103	16	US-10-437-963-115402	Sequence 115402, A
37	44	47.3	403	14	US-10-301-997-83	Sequence 83, Appl
38	44	47.3	568	9	US-09-815-242-5492	Sequence 5492, Ap
39	44	47.3	568	9	US-09-815-242-12629	Sequence 12629, A
40	44	47.3	568	12	US-10-282-122A-44222	Sequence 44222, A
41	43	46.2	71	16	US-10-437-963-115402	Sequence 115402, A
42	43	46.2	89	14	US-10-058-053A-62	Sequence 62, Appl
43	43	46.2	98	12	US-10-424-599-193800	Sequence 193800, A
44	43	46.2	119	12	US-10-257-174-41	Sequence 41, Appl
45	43	46.2	147	9	US-09-873-134-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-10-011-095-21
Sequence 21, Application US/10011095
Publication No. US20030045682A1
GENERAL INFORMATION:
APPLICANT: Afar, Daniel
APPLICANT: Hubert, Rene S.
APPLICANT: Leong, Kahan
APPLICANT: Raitano, Arthur B.
APPLICANT: Saittan, Douglas C.
APPLICANT: Mitchell, Steve Chappel
TITLE OF INVENTION: ANTIBODIES IMMUNOSPECIFIC FOR STEAP1 (AS AMENDED)
FILE REFERENCE: 511582001610
CURRENT APPLICATION NUMBER: US/10/011,095
CURRENT FILING DATE: 2001-12-06
PRIOR APPLICATION NUMBER: 09/523,873
PRIOR FILING DATE: 1999-06-01
PRIOR APPLICATION NUMBER: 60/087,520
PRIOR FILING DATE: 1998-06-01
PRIOR APPLICATION NUMBER: 60/091,183
PRIOR FILING DATE: 1998-06-30
NUMBER OF SEQ ID NOS: 32
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 21
LENGTH: 15
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: STEAP-1 PEPTIDE
US-10-011-095-21
Query Match 100.0%; Score 93; DB 14; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 WIDIKQFWYTPPTF 15
DB 1 WIDIKQFWYTPPTF 15

```
RESULT 2
US-10-010-667a-21
; Sequence 21, Application US/10010667a
; Publication No. US20030055217A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Hubert, Rene S.
; APPLICANT: Leong, Kahan
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Saffran, Douglas C.
; APPLICANT: Mitchell, Steve Chappell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; FILE REFERENCE: 51182001601
; CURRENT APPLICATION NUMBER: US/10/010,667a
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: 09/323,873
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: STEAP-1 PEPTIDE
US-10-010-667a-21

Query Match      100.0%; Score 93; DB 14; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WIDIKQFWVYTPPTF 15
DB      1 WIDIKQFWVYTPPTF 15

RESULT 3
US-10-165-044-39
; Sequence 39, Application US/10165044
; Publication No. US20030149531A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Rene S. Hubert
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas Saffran
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Steven Chappell Mitchell
; APPLICANT: Mary Faris
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; FILE REFERENCE: 51182-20016.02
; CURRENT APPLICATION NUMBER: US/10/165,044
; CURRENT FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: US 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: US 60/091,183
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 09/323,873
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: US 09/455,486
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: WO 99/62941
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/33040
; PRIOR FILING DATE: 2000-12-06
```

```
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-165-044-39

Query Match      100.0%; Score 93; DB 14; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WIDIKQFWVYTPPTF 15
DB      1 WIDIKQFWVYTPPTF 15

RESULT 4
US-09-747-835a-50
; Sequence 50, Application US/09747835a
; Patent No. US20020146692A1
; GENERAL INFORMATION:
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Asundi, Vinod
; APPLICANT: Drmanac, Radoje T
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIKE
; FILE REFERENCE: HYS-37C1P
; CURRENT APPLICATION NUMBER: US/09/747,835A
; CURRENT FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: US 09/729,739
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: US 09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 50
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-747-835a-50

Query Match      100.0%; Score 93; DB 9; Length 267;
Best Local Similarity 100.0%; Pred. No. 3.3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WIDIKQFWVYTPPTF 15
DB      211 WIDIKQFWVYTPPTF 225

RESULT 5
US-09-747-835a-51
; Sequence 51, Application US/09747835a
; Patent No. US20020146692A1
; GENERAL INFORMATION:
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
```

```
/ APPLICANT: Zhou, Ping
/ APPLICANT: Wang, Dunrui
/ APPLICANT: Zhang, Jie
/ APPLICANT: Ren, Feiyan
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Dirmnac, Radoje T
/ TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIKE
/ FILE REFERENCE: HXS-37CIP
/ CURRENT APPLICATION NUMBER: US/09/747,835A
/ PRIOR FILING DATE: 2002-03-08
/ PRIOR APPLICATION NUMBER: US 09/729,739
/ PRIOR FILING DATE: 2000-12-04
/ PRIOR APPLICATION NUMBER: US 09/653,450
/ PRIOR FILING DATE: 2000-08-31
/ PRIOR APPLICATION NUMBER: US 09/620,312
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: US 09/598,042
/ PRIOR FILING DATE: 2000-06-20
/ PRIOR APPLICATION NUMBER: US 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: US 09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 51
/ LENGTH: 267
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-747-835A-51

Query Match          100.0%; Score 93; DB 9; Length 267;
Best Local Similarity 100.0%; Pred. No. 3.3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WIDIKQFWYTPPTF 15
DB      211 WIDIKQFWYTPPTF 225

RESULT 6
US-10-312-312-50
/ Sequence 50, Application US/10312312
/ Publication No. US20040068097A1
/ GENERAL INFORMATION:
/ APPLICANT: Hyseq, Inc.
/ APPLICANT: Yamazaki, Victoria
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Zhou, Ping
/ APPLICANT: Wang, Dunrui
/ APPLICANT: Zhang, Jie
/ APPLICANT: Ren, Feiyan
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Dirmnac, Radoje T
/ TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIKE
/ FILE REFERENCE: 21272-015-061/HXS-37CIP
/ CURRENT APPLICATION NUMBER: US/10/312,312
/ PRIOR FILING DATE: 2002-12-10
/ PRIOR APPLICATION NUMBER: US 09/729,739
/ PRIOR FILING DATE: 2000-12-04
/ PRIOR APPLICATION NUMBER: US 09/653,450
/ PRIOR FILING DATE: 2000-08-31
/ PRIOR APPLICATION NUMBER: US 09/620,312
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: US 09/598,042
/ PRIOR FILING DATE: 2000-06-20
/ PRIOR APPLICATION NUMBER: US 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: US 09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 63
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 51
/ LENGTH: 267
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-312-312-51
```

```
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 50
/ LENGTH: 267
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-312-312-50

Query Match          100.0%; Score 93; DB 12; Length 267;
Best Local Similarity 100.0%; Pred. No. 3.3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WIDIKQFWYTPPTF 15
DB      211 WIDIKQFWYTPPTF 225

RESULT 7
US-10-312-312-51
/ Sequence 51, Application US/10312312
/ Publication No. US20040068097A1
/ GENERAL INFORMATION:
/ APPLICANT: Hyseq, Inc.
/ APPLICANT: Yamazaki, Victoria
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Zhou, Ping
/ APPLICANT: Wang, Dunrui
/ APPLICANT: Zhang, Jie
/ APPLICANT: Ren, Feiyan
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Dirmnac, Radoje T
/ TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIKE
/ FILE REFERENCE: 21272-015-061/HXS-37CIP
/ CURRENT APPLICATION NUMBER: US/10/312,312
/ PRIOR FILING DATE: 2002-12-10
/ PRIOR APPLICATION NUMBER: US 09/729,739
/ PRIOR FILING DATE: 2000-12-04
/ PRIOR APPLICATION NUMBER: US 09/653,450
/ PRIOR FILING DATE: 2000-08-31
/ PRIOR APPLICATION NUMBER: US 09/620,312
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: US 09/598,042
/ PRIOR FILING DATE: 2000-06-20
/ PRIOR APPLICATION NUMBER: US 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: US 09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 63
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 51
/ LENGTH: 267
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-312-312-51

Query Match          100.0%; Score 93; DB 12; Length 267;
Best Local Similarity 100.0%; Pred. No. 3.3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WIDIKQFWYTPPTF 15
DB      211 WIDIKQFWYTPPTF 225

RESULT 8
US-09-759-143-879
/ Sequence 879, Application US/09759143
/ Patent No. US20020022248A1
/ GENERAL INFORMATION:
/ APPLICANT: Xu, Jiangchun
/ APPLICANT: Dillon, Davin C.
/ APPLICANT: Mitcham, Jennifer L.
```

```

; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqun
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolck, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darriick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C23
; CURRENT APPLICATION NUMBER: US/09/759,143
; CURRENT FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-759-143-879
```

```

Query Match      100.0%; Score 93; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 4.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 WIDIKQFWYTPPTF 15
        |||
Db      277 WIDIKQFWYTPPTF 291
```

```

RESULT 9
US-09-780-669-879
; Sequence 879, Application US/09780669
; Patent No. US20020051977A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqun
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolck, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darriick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C24
; CURRENT APPLICATION NUMBER: US/09/780,669
; CURRENT FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 943
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-780-669-879
```

```

Query Match      100.0%; Score 93; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 4.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 WIDIKQFWYTPPTF 15
        |||
Db      277 WIDIKQFWYTPPTF 291
```

```

RESULT 10
US-09-822-827-879
; Sequence 879, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C1
; CURRENT APPLICATION NUMBER: US/09/822,827
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-822-827-879
```

```

Query Match      100.0%; Score 93; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 4.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 WIDIKQFWYTPPTF 15
        |||
Db      277 WIDIKQFWYTPPTF 291
```

```

RESULT 11
US-09-802-520-11
; Sequence 11, Application US/09802520
; Publication No. US20020187472A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Paris, Mary
; APPLICANT: Chen, Hwei-Mei
; APPLICANT: Ison, Craig H.
; TITLE OF INVENTION: STEAP-RELATED PROTEIN
; FILE REFERENCE: PC-0037 US
; CURRENT APPLICATION NUMBER: US/09/802,520
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PERL Program
; SEQ ID NO 11
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020187472A1 96572948
US-09-802-520-11
```

```

Query Match      100.0%; Score 93; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 4.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 WIDIKQFWYTPPTF 15
        |||
Db      277 WIDIKQFWYTPPTF 291
```

```

RESULT 12
US-09-895-793-879
```



```

; Sequence 879, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiaqichun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiaq, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Derrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aljun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hurai, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.534C2
; CURRENT APPLICATION NUMBER: US/09/895,793
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-895-793-879

Query Match      100.0%; Score 93; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 4.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WIDIKQFVWYTPPTF 15
DB      277 WIDIKQFVWYTPPTF 291

RESULT 13
US-09-895-814-879
; Sequence 879, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiaqichun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiaq, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Derrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aljun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hurai, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa

```

```

; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C26
; CURRENT APPLICATION NUMBER: US/09/895,814
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-895-814-879

Query Match      100.0%; Score 93; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 4.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WIDIKQFVWYTPPTF 15
DB      277 WIDIKQFVWYTPPTF 291

RESULT 14
US-10-205-267-13
; Sequence 13, Application US/10205267
; Publication No. US20030064397A1
; GENERAL INFORMATION:
; APPLICANT: Spence, Kimberly M.
; APPLICANT: Rickett, Paula K.
; APPLICANT: Lal, Preeti G.
; APPLICANT: Ison, Craig H.
; TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN PROSTATE AND
; TITLE OF INVENTION: TUMORS
; FILE REFERENCE: PV-0008 CIP
; CURRENT APPLICATION NUMBER: US/10/205,267
; CURRENT FILING DATE: 2002-07-24
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PERL Program
; SEQ ID NO 13
; LENGTH: 339
; TYPE: PRT
; ORGANISM:
; PUBLICATON INFORMATION:
; DATABASE ACCESSION NUMBER: Genbank ID No: 96572948
; US-10-205-267-13

Query Match      100.0%; Score 93; DB 12; Length 339;
Best Local Similarity 100.0%; Pred. No. 4.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WIDIKQFVWYTPPTF 15
DB      277 WIDIKQFVWYTPPTF 291

RESULT 15
US-10-408-009-2
; Sequence 2, Application US/10408009
; Publication No. US20040072196A1
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Stephen C. Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USES THEREOF
; FILE REFERENCE: 511582001603
; CURRENT APPLICATION NUMBER: US/10/408,009
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 09/455,486
; PRIOR FILING DATE: 1999-12-06

```

```

; PRIOR APPLICATION NUMBER: 09/323,873
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-009-2

```

```

Query Match      100.0%; Score 93; DB 12; Length 339;
Best Local Similarity 100.0%; Pred. No. 4.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 WIDIKQFWYTPPTF 15
          |||||
Db      277 WIDIKQFWYTPPTF 291

```

Search completed: July 12, 2004, 14:27:49
 Job time : 26.5116 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 14:04:08 ; Search time 9.24419 Seconds

(Without alignments)
83.770 Million cell updates/sec

Title: US-10-010-667a-2_COPY_277_291
Sequence: 93
1 WIDIKQFWVYTPPTF 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 segs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: /cgn2_6/prodata/2/1aa/5A COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCTUS COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	93	100.0	15	4	US-09-323-873A-21
2	93	100.0	339	4	US-09-323-873A-2
3	93	100.0	339	4	US-09-323-873A-2
4	93	100.0	268	3	US-08-965-056-13
5	45	48.4	269	3	US-08-965-056-13
6	45	48.4	353	1	US-08-118-270-45
7	45	48.4	353	5	PCT-US93-08528-45
8	45	48.4	489	4	US-09-134-000C-4888
9	45	48.4	854	4	US-09-206-551-17
10	45	48.4	865	3	US-07-956-483-13
11	45	48.4	887	3	US-08-472-240A-4
12	44	47.3	403	4	US-08-887-534A-83
13	44	47.3	403	4	US-09-527-431-83
14	43	46.2	351	4	US-08-817-441-47
15	43	46.2	862	4	US-09-206-551-15
16	43	46.2	873	2	US-08-912-129A-61
17	43	46.2	877	4	US-08-817-441-102
18	42	45.2	268	3	US-08-965-056-26
19	42	45.2	268	3	US-08-965-056-68
20	42	45.2	268	3	US-08-965-056-71
21	42	45.2	446	4	US-09-543-681A-5890
22	42	45.2	533	4	US-09-107-532A-7241
23	41.5	44.6	231	3	US-08-448-489-19
24	41	44.1	237	3	US-08-388-353-642
25	41	44.1	237	3	US-08-488-551B-642
26	41	44.1	243	3	US-08-965-056-73
27	41	44.1	268	3	US-08-965-056-8

28	41	44.1	268	3	US-08-965-056-13	Sequence 13, Appl
29	41	44.1	268	3	US-08-965-056-14	Sequence 14, Appl
30	41	44.1	268	3	US-08-965-056-70	Sequence 70, Appl
31	41	44.1	268	3	US-08-965-056-72	Sequence 72, Appl
32	41	44.1	269	3	US-08-965-056-21	Sequence 21, Appl
33	41	44.1	269	3	US-08-965-056-22	Sequence 22, Appl
34	41	44.1	269	3	US-08-965-056-25	Sequence 25, Appl
35	41	44.1	269	3	US-08-965-056-37	Sequence 37, Appl
36	41	44.1	270	3	US-08-965-056-67	Sequence 46, Appl
37	41	44.1	351	1	US-08-470-202-46	Sequence 46, Appl
38	41	44.1	351	1	US-08-471-770-46	Sequence 46, Appl
39	41	44.1	351	2	US-08-468-059-46	Sequence 46, Appl
40	41	44.1	351	3	US-09-109-916-46	Sequence 46, Appl
41	41	44.1	351	4	US-09-886-156-46	Sequence 46, Appl
42	41	44.1	351	4	US-09-886-149-46	Sequence 46, Appl
43	41	44.1	351	4	US-09-886-150-46	Sequence 46, Appl
44	41	44.1	351	4	US-09-886-159-46	Sequence 46, Appl
45	41	44.1	602	3	US-09-257-490-15	Sequence 15, Appl

ALIGNMENTS

```
RESULT 1
US-09-323-873A-21
; Sequence 21, Application US/09323873A
; Patent No. 6329503
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Steve Chappell Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USES THEREOF
; FILE REFERENCE: 129.16TSU2
; CURRENT APPLICATION NUMBER: US/09/323,873A
; CURRENT FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: STEAP-1 PREPTIDE
US-09-323-873A-21

Query Match      100.0%; Score 93; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.5e+08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WIDIKQFWVYTPPTF 15
Db      1 WIDIKQFWVYTPPTF 15

RESULT 2
US-09-323-873A-2
; Sequence 2, Application US/09323873A
; Patent No. 6329503
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Steve Chappell Mitchell
```

```
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; FILE REFERENCE: 129.160SU2
; CURRENT APPLICATION NUMBER: US/09/323,873A
; PRIOR FILING DATE: 1999-06-01
; PRIOR FILING DATE: 1998-06-01
; PRIOR FILING DATE: 1998-06-30
; PRIOR FILING DATE: 1998-06-30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo Sapiens
; US-09-323-873A-2

Query Match      100.0%; Score 93; DB 4; Length 339;
Best Local Similarity 100.0%; Pred. No. 8.3e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  WIDIKQFWWYTPPTF 15
Db      277 WIDIKQFWWYTPPTF 291

RESULT 3
US-09-685-166A-879
; Sequence 879, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
; APPLICANT: Xu, Jianshun
; APPLICANT: Dillon, David C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685,166A
; CURRENT FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-685-166A-879

Query Match      100.0%; Score 93; DB 4; Length 339;
Best Local Similarity 100.0%; Pred. No. 8.3e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  WIDIKQFWWYTPPTF 15
Db      277 WIDIKQFWWYTPPTF 291

RESULT 4
US-08-965-056-39
; Sequence 39, Application US/08965056
```

```
; Patent No. 6271198
; GENERAL INFORMATION:
; APPLICANT: Andrew C. Braisted
; APPLICANT: J. Kevin Judice
; APPLICANT: Robert S. McDowell
; APPLICANT: J. Christopher Phelan
; APPLICANT: Melissa A. Starovasnik
; APPLICANT: James A. Wells
; TITLE OF INVENTION: Constrained Helical Peptides and Methods of
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESS: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/965,056
; FILING DATE: 05-NO. 6271198-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, Ph.D., Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: P1005R2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8674
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 268 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-08-965-056-39

Query Match      48.4%; Score 45; DB 3; Length 268;
Best Local Similarity 60.0%; Pred. No. 17;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1  WIDIKQFWY 10
Db      202 WFDITQWLWY 211

RESULT 5
US-08-965-056-33
; Sequence 33, Application US/08965056
; Patent No. 6271198
; GENERAL INFORMATION:
; APPLICANT: Andrew C. Braisted
; APPLICANT: J. Kevin Judice
; APPLICANT: Robert S. McDowell
; APPLICANT: J. Christopher Phelan
; APPLICANT: Melissa A. Starovasnik
; APPLICANT: James A. Wells
; TITLE OF INVENTION: Constrained Helical Peptides and Methods of
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESS: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/965,056
FILING DATE: 05-NO. 6271198-1997
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Torchia, PhD., Timothy E.
REGISTRATION NUMBER: 36,700
REFERENCE/DOCKET NUMBER: P1005R2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-8674
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 269 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-965-056-33

Query Match 48.4%; Score 45; DB 3; Length 269;
Best Local Similarity 60.0%; Pred. No. 17;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 WIDIKQFVWY 10
DB 203 WFDITQWLV 212

RESULT 6
US-08-118-270-45
Sequence 45, Application US/08118270
Patent No. 5508384
GENERAL INFORMATION:
APPLICANT: Murphy, Randall B.
ATTORNEY/AGENT INFORMATION:
NAME: Schuster, David I.
TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
NUMBER OF SEQUENCES: 348
CORRESPONDENCE ADDRESS:
ADDRESSER: BROWDY AND NEWMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/118,270
FILING DATE: 09-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/943,236
FILING DATE: 10-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Townsend, Kevin G.
REGISTRATION NUMBER: 34,033
REFERENCE/DOCKET NUMBER: MURPHY=2A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 353 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide

US-08-118-270-45

Query Match 48.4%; Score 45; DB 1; Length 353;
Best Local Similarity 33.3%; Pred. No. 22;
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 WIDIKQFVWYPTPTF 15
DB 56 WVELYNFIMHPWAF 70

RESULT 7
PCT-US93-08528-45
Sequence 45, Application PC/TUS9308528
GENERAL INFORMATION:
APPLICANT: New York University
TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
NUMBER OF SEQUENCES: 348
CORRESPONDENCE ADDRESS:
ADDRESSER: BROWDY AND NEWMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/08528
FILING DATE: 09-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/943,236
FILING DATE: 10-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Townsend, Kevin G.
REGISTRATION NUMBER: 34,033
REFERENCE/DOCKET NUMBER: MURPHY=2 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 353 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US93-08528-45

Query Match 48.4%; Score 45; DB 5; Length 353;
Best Local Similarity 33.3%; Pred. No. 22;
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 WIDIKQFVWYPTPTF 15
DB 56 WVELYNFIMHPWAF 70

RESULT 8
US-09-134-000C-4888
Sequence 4888, Application US/09134000C
Patent No. 6617156
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 03796-0032
CURRENT APPLICATION NUMBER: US/09/134,000C

```

; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4888
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; US-09-134-000C-4888

Query Match
Best Local Similarity 48.4%; Score 45; DB 4; Length 489;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 WIDIKQFVWYTP 13
Db 450 WPEITQIWLAP 462

RESULT 9
US-09-206-551-17
; Sequence 17, Application US/09206551B
; Patent No. 6521739
; GENERAL INFORMATION:
; APPLICANT: Hahn, Beatrice H.
; APPLICANT: Gao, Feng
; APPLICANT: Marx, Preston A.
; APPLICANT: Shaw, George M.
; APPLICANT: Smith, Stephen M.
; APPLICANT: Georges-Courbot, Marie Claude
; APPLICANT: Lu, Chang Yong
; TITLE OF INVENTION: Complete Genome Sequences of a Simian
; TITLE OF INVENTION: Immunodeficiency Virus from a Red-Capped
; FILE REFERENCE: D6286
; CURRENT APPLICATION NUMBER: US/09/206,551B
; CURRENT FILING DATE: 1998-12-07
; NUMBER OF SEQ ID NOS: 58
; SEQ ID NO 17
; LENGTH: 854
; TYPE: PRT
; ORGANISM: Simian immunodeficiency virus
; FEATURE:
; OTHER INFORMATION: Amino acid sequence of homologous region of
; US-09-206-551-17

Query Match
Best Local Similarity 48.4%; Score 45; DB 4; Length 854;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 WIDIKQFVWY 10
Db 670 WPDITQWLWY 679

RESULT 10
US-07-956-483-13
; Sequence 13, Application US/07956483
; Patent No. 6261799
; GENERAL INFORMATION:
; APPLICANT: KIENY, Marie-Paule
; TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/956,483
; FILING DATE: 31-DEC-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 92/19742
; FILING DATE: 12-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91 05392
; FILING DATE: 02-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Crane-Feury, Sharon E
; REGISTRATION NUMBER: 36,113
; REFERENCE/DOCKET NUMBER: 017753-005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 865 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-07-956-483-13
```

```

Query Match
Best Local Similarity 48.4%; Score 45; DB 3; Length 865;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```

QY 1 WIDIKQFVWY 10
Db 681 WPDITQWLWY 690

RESULT 11
US-08-472-240A-4
; Sequence 4, Application US/08472240A
; Patent No. 6284248
; GENERAL INFORMATION:
; APPLICANT: KIENY, Marie-Paule
; TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,240A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/956,483
; FILING DATE: 31-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Teskin, Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 017753-055
; TELECOMMUNICATION INFORMATION:
```

TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 887 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Protein
LOCATION: 1..854
US-08-472-240A-4

Query Match 48.4%; Score 45; DB 3; Length 887;
Best Local Similarity 60.0%; Pred. No. 56;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MIDIKOFWY 10
DB 703 MFDITQWLMY 712

RESULT 12
US-08-887-534A-83
Sequence 83, Application US/08887534A
Patent No. 6455323
GENERAL INFORMATION:
ATTORNEY/AGENT INFORMATION:
NAME: Holden, David W.
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 28341/33996
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-6300
TELEX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 403 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-887-534A-83

Query Match 47.3%; Score 44; DB 4; Length 403;
Best Local Similarity 58.3%; Pred. No. 36;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 DIKQFWYTPPT 14
DB 305 DFKVTIWTAPT 316

RESULT 13

US-09-527-431-83
Sequence 83, Application US/09527431
Patent No. 6485899
GENERAL INFORMATION:
ATTORNEY/AGENT INFORMATION:
NAME: Holden, David W.
REGISTRATION NUMBER: 106
REFERENCE/DOCKET NUMBER: 28341/33996
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-6300
TELEX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 403 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-527-431-83

Query Match 47.3%; Score 44; DB 4; Length 403;
Best Local Similarity 58.3%; Pred. No. 36;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 DIKQFWYTPPT 14
DB 305 DFKVTIWTAPT 316

RESULT 14
US-08-817-441-47
Sequence 47, Application US/08817441
Patent No. 639294
GENERAL INFORMATION:
ATTORNEY/AGENT INFORMATION:
NAME: CHARBEAU, PIERRE
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 28341/33996
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-6300
TELEX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 403 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-817-441-47

Query Match 47.3%; Score 44; DB 4; Length 403;
Best Local Similarity 58.3%; Pred. No. 36;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 DIKQFWYTPPT 14
DB 305 DFKVTIWTAPT 316

COUNTRY: USA
 ZIP: 20005-3315
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/817,441
 FILING DATE: 11-JUL-1997
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/FR 95/01391
 FILING DATE: 20-OCT-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: FR 9412554
 FILING DATE: 20-OCT-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: FR 9502526
 FILING DATE: 03-MAR-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Meyers, Kenneth J.
 REGISTRATION NUMBER: 25,146
 REFERENCE/DOCKET NUMBER: 03260.6005-00000
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-408-4000
 TELEFAX: 202-408-4400
 INFORMATION FOR SEQ ID NO: 47:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 351 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-817-441-47

Query Match 46.2%; Score 43; DB 4; Length 351;
 Best Local Similarity 50.0%; Pred. No. 45;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 WIDIKQFWY 10
 DB 160 WIDITKLMY 169

RESULT 15
 US-09-206-551-15
 ; Sequence 15, Application US/09206551B
 ; Patent No. 6521739
 ; GENERAL INFORMATION:
 ; APPLICANT: Hahn, Beatrice H.
 ; APPLICANT: Gao, Feng
 ; APPLICANT: Marx, Preston A.
 ; APPLICANT: Shaw, George M.
 ; APPLICANT: Smith, Stephen M.
 ; APPLICANT: Georges-Courbot, Marie Claude
 ; APPLICANT: Lu, Chang Yong
 ; TITLE OF INVENTION: Complete Genome Sequences of a Simian
 ; TITLE OF INVENTION: Immunodeficiency Virus from a Red-Capped
 ; FILE REFERENCE: D6286
 ; CURRENT APPLICATION NUMBER: US/09/206,551B
 ; CURRENT FILING DATE: 1998-12-07
 ; NUMBER OF SEQ ID NOS: 58
 ; SEQ ID NO 15
 ; LENGTH: 862
 ; TYPE: PRT
 ; ORGANISM: Simian immunodeficiency virus
 ; FEATURE:
 ; OTHER INFORMATION: Amino acid sequence of homologous region of
 ; OTHER INFORMATION: O_ANT10 lentiviral env protein
 US-09-206-551-15

Query Match 46.2%; Score 43; DB 4; Length 862;

Best Local Similarity 50.0%; Pred. No. 1.1e+02;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 QY 1 WIDIKQFWY 10
 DB 671 WIDITKLMY 680

Search completed: July 12, 2004, 14:12:24
 Job time : 9.24419 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 13:51:07 ; Search time 32.9651 Seconds

(without alignments)
128.567 Million cell updates/sec

Title: US-10-010-667a-2_COPY_277_291

Sequence: 1 WIDIKQFWYTPPTF 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003s:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	93	100.0	15	3	AAy58200 Human STR
2	93	100.0	15	4	AAE02788 Extracell
3	93	100.0	267	6	ABU60886 Human G P
4	93	100.0	267	6	ABU60887 Human G P
5	93	100.0	339	3	AAy58194 Human STR
6	93	100.0	339	4	AAm01282 P789P ami
7	93	100.0	339	4	AAU69927 Human pro
8	93	100.0	339	4	AAm78845 Human pro
9	93	100.0	339	4	AAU71818 Prostate
10	93	100.0	339	5	ABB95387 Human P78
11	93	100.0	339	5	ABG61813 Prostate
12	93	100.0	339	6	ABU98383 Novel hum
13	93	100.0	339	6	ABU98414 STRAP-1 v
14	93	100.0	339	6	ABU98430 STRAP-1 v
15	93	100.0	339	6	ABU98430 STRAP-1 v
16	93	100.0	339	7	ABU63313 Prostate
17	93	100.0	339	7	ABU63313 Human six
18	93	100.0	339	7	ABU63313 Human six
19	93	100.0	339	7	ABU63313 Human six
20	93	100.0	339	7	ABU63313 Human six
21	93	100.0	339	7	ABU63313 Human six
22	93	100.0	339	7	ABU63313 Human six
23	93	100.0	339	7	ABU63313 Human six
24	93	100.0	339	7	ABU63313 Human six
25	93	100.0	339	7	ABU63313 Human six

26	45	48.4	269	2	AAy22837 SEQ ID NO
27	45	48.4	269	5	ABG68308 Envelope
28	45	48.4	269	5	ABU57715 Human Imm
29	45	48.4	353	2	AAr48724 G-Protein
30	45	48.4	353	2	AAW02696 G-Protein
31	45	48.4	483	6	ABU29474 Protein e
32	45	48.4	846	3	ABG69345 HIV-1 non
33	45	48.4	861	2	AAW43074 HIV-1 gpl
34	45	48.4	861	2	AAW43074 HIV-1 gpl
35	45	48.4	871	6	AAW02696 HIV prote
36	45	48.4	871	6	AAW02696 HIV prote
37	45	48.4	871	6	AAW02696 HIV prote
38	45	48.4	871	6	AAW02696 HIV prote
39	45	48.4	871	6	AAW02696 HIV prote
40	45	48.4	871	6	AAW02696 HIV prote
41	45	48.4	871	6	AAW02696 HIV prote
42	45	48.4	871	6	AAW02696 HIV prote
43	45	48.4	871	6	AAW02696 HIV prote
44	45	48.4	871	6	AAW02696 HIV prote
45	45	48.4	871	6	AAW02696 HIV prote

ALIGNMENTS

RESULT 1	AAy58200	standard; peptide; 15 AA.
ID	AAy58200	standard; peptide; 15 AA.
XX	AAy58200	standard; peptide; 15 AA.
AC	AAy58200	standard; peptide; 15 AA.
XX	AAy58200	standard; peptide; 15 AA.
DT	14-MAR-2000	(first entry)
XX	14-MAR-2000	(first entry)
DE	Human STRAP-1 peptide, corresponding to STRAP-1 extracellular region 3.	
XX	Human STRAP-1 peptide, corresponding to STRAP-1 extracellular region 3.	
KW	Serpinin transmembrane antigen of the prostate; STRAP-1; prostate;	
KW	transmembrane domain; type IIIa membrane protein; expression; cancer;	
KW	prostate cancer; bladder cancer; colon cancer; pancreatic cancer;	
KW	ovarian cancer; tumour antigen; immunisation; immune response; cellular;	
KW	humoral; anticancer vaccine; antibody; detection; diagnosis; prognosis;	
KW	monitoring; susceptibility; therapeutic inhibitor; drug targeting;	
XX	recombinant protein.	
OS	Synthetic.	
XX	Homo sapiens.	
XX	Homo sapiens.	
PN	WO962941-A2.	
XX	WO962941-A2.	
PD	09-DEC-1999.	
XX	09-DEC-1999.	
PF	01-JUN-1999;	99WO-US012157.
XX	01-JUN-1999;	99WO-US012157.
PR	01-JUN-1998;	98US-0087520P.
PR	01-JUN-1998;	98US-0087520P.
XX	01-JUN-1998;	98US-0087520P.
PA	(UOEG-) UROGENESYS INC.	
PA	(AFAR/) AFAR D E.	
PA	(HUBB/) HUBB R S.	
PA	(LEON/) LEON K.	
PA	(RAIT/) RAITANO A B.	
PA	(SAFF/) SAFFRAN D C.	
PI	Afar DE, Hubert RS, Leong K, Raitano AB, Saffran DC;	
DR	WPI, 2000-072832/06.	
XX	WPI, 2000-072832/06.	
PT	Novel proteins useful as diagnostic markers and therapeutic targets,	
PT	particularly for prostatic cancer.	
XX	Novel proteins useful as diagnostic markers and therapeutic targets,	
XX	particularly for prostatic cancer.	
PS	Disclosure; Page 22; 83pp; English.	
CC	Sequences AAy58198-Y58200 represent synthetic peptides that correspond to	
CC	the extracellular regions of STRAP-1 (serpinin transmembrane antigen of	

the prostate, AAV58194). These peptides were used to raise monoclonal anti-STRAP-1 antibodies. STRAP-1 is the prototype member of the STRAP family of proteins (AAV58194-Y58197) which exhibit a high degree of structural conservation, but which show no significant structural homology to known human proteins. The STRAP-1 gene has been localised to chromosome 7p22. STRAP-1 is thought to be a type IIIa membrane protein and is expressed predominantly in prostate cells in normal human tissues. Structurally, STRAP-1 is a 339 amino acid protein characterised by six transmembrane domains and intracellular N- and C-termini, suggesting that it folds in a "serpentine" manner into three extracellular and two intracellular loops. STRAP-1 mRNA and protein expression is maintained at high levels and throughout all stages of prostate cancer. STRAP-1 mRNA and/or protein is also overexpressed in certain other cancers, including bladder, colon, pancreatic and ovarian cancer. The function of the STRAP proteins is not known. They may be ion channels (from the presence of six transmembrane domains, a feature which is shared by certain ion channels) or gap-junction proteins (from immunohistochemical staining). STRAP-1 and STRAP-2 are cell-surface and humoral immune responses against STRAP-1 protein induces cellular and humoral immune responses against STRAP-1 expressing cells. STRAP proteins may be used to identify specific-binding agents, to produce anticancer vaccines and to generate specific antibodies. The antibodies may be used for detection, prognosis, and monitoring of cancers (or susceptibility to cancer), as therapeutic inhibitors or to target therapeutic agents to their site of action. STRAP nucleic acids may be used for recombinant protein production, as diagnostic and prognostic reagents, for identifying STRAP-expressing cells for screening inhibitors of STRAP expression and for therapeutic modulation/inhibition of STRAP expression. Since high levels of STRAP proteins are exposed on the cell surface, they are easily targeted by systemically administered agents, and because they are expressed mainly on prostatic epithelial cells, agents targeted to them should have minimal side effects on other tissues

SQ Sequence 15 AA;

Query Match 100.0%; Score 93; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 6.5e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WIDIKQFWWYTPPTF 15
| | | | | | | | | | | | | | | | |
DB 1 WIDIKQFWWYTPPTF 15

RESULT 2
AAE02788 standard; peptide; 15 AA.
ID AAE02788
XX
AC AAE02788;
XX
DT 06-AUG-2001 (first entry)

DE Extracellular loop #3 of human STRAP-1, suitable for cloning into pFc.
XX
XX Human, cytosolic; antiproliferative; vaccine; gene therapy;
KW six transmembrane epithelial antigen of the prostate-1; STRAP-1;
KM chromosome 7p22.3; cancer; prostate; colon; bladder; pancreatic; ovarian;
KW lung; extracellular loop; serpentine transmembrane antigen.

XX Homo sapiens.

XX MO200140276-A2.

XX 07-JUN-2001.

XX 06-DEC-2000; 2000WO-US033040.

XX 06-DEC-1999; 99US-00455486.

XX (UROG-) UROGENESYS INC.

PA Afar DEH, Hubert RS, Raitano AB, Saffran DC, Mitchell SC;
PI Paris M, Jakobovits A;

XX WPI; 2001-367804/38.

DR New STRAP (six transmembrane epithelial antigen of the prostate)
XX proteins, expressed in human cancers, useful for detecting and treating
PT cancer.
XX

XX Example 19; Page 102; 187pp; English.

XX The present invention relates to human six transmembrane epithelial
CC antigen of the prostate (STRAP) protein. STRAP is a member of cell
CC surface serpentine transmembrane antigens. STRAP gene is used in gene
CC therapy, inhibiting the development or progression of a cancer (eg.
CC prostate, colon, bladder, lung, ovarian and pancreatic) expressing STRAP
CC or inhibiting growth or killing cells expressing STRAP in a patient,
CC comprising administering a vaccine composition to the patient. Treating a
CC patient with a cancer that expresses STRAP, or inhibiting growth or
CC killing cells expressing STRAP, comprises administering to the patient a
CC vector encoding single chain monoclonal antibody that comprises the
CC variable domains of the heavy and light chains of the monoclonal antibody
CC that specifically binds to STRAP, such that the vector delivers the
CC single chain monoclonal antibody coding sequence to the cancer cells and
CC the encoded single chain monoclonal antibody is expressed
CC intracellularly. The present sequence is extracellular loop of STRAP-1
CC suitable for cloning into pFc, which is used in the invention. STRAP-1
CC gene is located on chromosome 7p22.3

SQ Sequence 15 AA;

Query Match 100.0%; Score 93; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 6.5e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WIDIKQFWWYTPPTF 15
| | | | | | | | | | | | | | | | |
DB 1 WIDIKQFWWYTPPTF 15

RESULT 3
ABU60886 standard; protein; 267 AA.
ID ABU60886
XX
AC ABU60886;
XX
DT 08-MAY-2003 (first entry)

DE Human G protein-coupled receptor-related (GPCR-related) protein #3.

XX Human; G protein-coupled receptor; GPCR; receptor; ophthalmic disease;
KW neurological disease; immunological disease; nephritic disease; cancer;
KW hormonal dysfunction; neoplasm; atherosclerosis; diabetes; chemotherapy;
KW anaemia; irradiation; erythroid precursor; erythroid cell; ulcer; organ;
KW pressure ulcer; vascular insufficiency; surgical wound; traumatic wound;
KW pancreas; liver; intestine; kidney; skin; endothelium; skeletal muscle;
KW smooth muscle; cardiac muscle; vascular tissue.

XX Homo sapiens.

XX US2002146692-A1.

XX 10-OCT-2002.

XX 21-DEC-2000; 2000US-00747835.

XX 21-JAN-2000; 2000US-00488725.

XX 25-APR-2000; 2000US-00552317.

XX 20-JUN-2000; 2000US-00598042.

XX 19-JUL-2000; 2000US-00620312.

XX 31-AUG-2000; 2000US-00653450.

XX 04-DEC-2000; 2000US-00729739.

PA (YAMA/) YAMAZAKI V.
PI (TANG/) TANG Y T.

PA (LIUC/) LIU C.
PA (ZHOU/) ZHOU P.
PA (WANG/) WANG D.
PA (ZHANG/) ZHANG J.
PA (ZHANG/) ZHANG J.
PA (REN/) REN F.
PA (ASUNDI/) ASUNDI V.
PA (DRMA/) DRMANAC R T.
PI Yamazaki V, Tang YT, Liu C, Zhou P, Wang D, Zhang J, Ren F,
PI Asundi V, Drmanac RT;
XX WPI; 2003-174164/17.
XX
XX
XX Isolated polynucleotide for treating, e.g. ophthalmic diseases, comprises
XX specified nucleotide sequences, or their translated protein coding
XX portion, mature protein coding portion, extracellular portion, or active
XX domain.
XX
XX Example 3; Fig 3; 158pp; English.
XX
XX The invention relates to human G protein-coupled receptor-like (GPCR-
XX like) proteins and the polynucleotides encoding them. The isolated
XX sequences are used for the treatment of diseases of ophthalmic,
XX neurological, immunological, or nephritic systems. They may also be used
XX to treat hormonal dysfunction, cancer and other neoplasia,
XX atherosclerosis, diabetes, in treating various anaemias or for use in
XX conjunction with irradiation/chemotherapy to stimulate the production of
XX erythroid precursors and/or erythroid cells. The sequences may be used to
XX promote better or faster closure of non-healing wounds, including
XX pressure ulcers, ulcers associated with vascular insufficiency and
XX surgical and traumatic wounds and in generation or regeneration of other
XX tissues, e.g. organs (e.g. pancreas, liver, intestine, kidney, skin,
XX endothelium), muscle (smooth, skeletal, or cardiac) and vascular tissue
XX or for promoting the growth of cells comprising such tissues. Sequences
XX ABU60884-ABU60895 represent human GPCR-related proteins of the invention
SQ Sequence 267 AA;
QY
Db 1 WIDIKQFWYTPPTF 15
211 WIDIKQFWYTPPTF 225
RESULT 4
ABU60887
ID ABU60887 standard; protein; 267 AA.
XX
XX ABU60887;
XX
XX 08-MAY-2003 (first entry)
XX
XX Human G protein-coupled receptor-related (GPCR-related) protein #4.
XX
XX Human; G protein-coupled receptor; GPCR; receptor; ophthalmic disease;
XX neurological disease; immunological disease; nephritic disease; cancer;
XX hormonal dysfunction; neoplasm; atherosclerosis; diabetes; chemotherapy;
XX anaemia; irradiation; erythroid precursor; erythroid cell; ulcer; organ;
XX pressure ulcer; vascular insufficiency; surgical wound; traumatic wound;
XX pancreas; liver; intestine; kidney; skin; endothelium; skeletal muscle;
XX smooth muscle; cardiac muscle; vascular tissue.
XX
XX Homo sapiens.
XX
XX US2002146692-A1.
XX
XX 10-OCT-2002.
XX
XX 21-DEC-2000; 2000US-00747835.
XX

PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 31-AUG-2000; 2000US-00653450.
PR 04-DEC-2000; 2000US-00729739.
XX
XX (YAMA/) YAMAZAKI V.
XX (TANG/) TANG Y T.
XX (LIUC/) LIU C.
XX (ZHOU/) ZHOU P.
XX (WANG/) WANG D.
XX (ZHANG/) ZHANG J.
XX (REN/) REN F.
XX (ASUNDI/) ASUNDI V.
XX (DRMA/) DRMANAC R T.
PI Yamazaki V, Tang YT, Liu C, Zhou P, Wang D, Zhang J, Ren F,
PI Asundi V, Drmanac RT;
XX WPI; 2003-174164/17.
XX
XX Isolated polynucleotide for treating, e.g. ophthalmic diseases, comprises
XX specified nucleotide sequences, or their translated protein coding
XX portion, mature protein coding portion, extracellular portion, or active
XX domain.
XX
XX Example 3; Fig 4; 158pp; English.
XX
XX The invention relates to human G protein-coupled receptor-like (GPCR-
XX like) proteins and the polynucleotides encoding them. The isolated
XX sequences are used for the treatment of diseases of ophthalmic,
XX neurological, immunological, or nephritic systems. They may also be used
XX to treat hormonal dysfunction, cancer and other neoplasia,
XX atherosclerosis, diabetes, in treating various anaemias or for use in
XX conjunction with irradiation/chemotherapy to stimulate the production of
XX erythroid precursors and/or erythroid cells. The sequences may be used to
XX promote better or faster closure of non-healing wounds, including
XX pressure ulcers, ulcers associated with vascular insufficiency and
XX surgical and traumatic wounds and in generation or regeneration of other
XX tissues, e.g. organs (e.g. pancreas, liver, intestine, kidney, skin,
XX endothelium), muscle (smooth, skeletal, or cardiac) and vascular tissue
XX or for promoting the growth of cells comprising such tissues. Sequences
XX ABU60884-ABU60895 represent human GPCR-related proteins of the invention
SQ Sequence 267 AA;
QY
Db 1 WIDIKQFWYTPPTF 15
211 WIDIKQFWYTPPTF 225
RESULT 5
AAV58194
ID AAV58194 standard; protein; 339 AA.
XX
XX AAV58194;
XX
XX 14-MAR-2000 (first entry)
XX
XX Human STRAP-1 protein.
XX
XX Serpentine transmembrane antigen of the prostate; STRAP-1; prostate;
XX transmembrane domain; type IIIa membrane protein; expression; cancer;
XX prostate cancer; bladder cancer; colon cancer; pancreatic cancer;
XX ovarian cancer; tumour antigen; immunisation; immune response; cellular;
XX humoral; anticancer vaccine; antibody; detection; diagnosis; prognosis;
XX monitoring; susceptibility; therapeutic inhibitor; drug targeting;
XX recombinant protein.
XX

Accession	Protein Name	Location/Qualifiers
XX OS	Homo sapiens.	
XX FH	Key	Location/Qualifiers
PT	Region	1..69
FT		/note= "Intracellular region 1"
FT		70..91
FT	Domain	/note= "Transmembrane domain 1"
FT		92..113
PT	Region	/note= "Extracellular region 1 (AAV58198) "
FT	Domain	114..136
FT		/note= "Transmembrane domain 2"
FT	Region	137..162
FT		/note= "Intracellular region 2"
FT		163..184
FT		/note= "Transmembrane domain 3"
FT	Region	185..218
FT		/note= "Extracellular region 2 (AAV58199) "
FT	Domain	219..241
FT		/note= "Transmembrane domain 4"
FT		242..253
FT	Region	/note= "Intracellular region 3"
FT		252..276
FT	Domain	/note= "Transmembrane domain 5"
FT		277..291
FT	Region	/note= "Extracellular region 3 (AAV58200) "
FT	Domain	292..313
FT		/note= "Transmembrane domain 6"
FT	Region	314..339
FT		/note= "Intracellular region 4"
XX PN		
XX	W09362941-A2.	
XX PD	09-DEC-1999.	
XX		
PF	01-JUN-1999;	99WC-US012157.
XX		
PR	01-JUN-1998;	98US-0087520P.
PR	30-JUN-1998;	98US-0091183P.
XX		
PA	(UROG-) UROGENESYS INC.	
PA	(AFAR/) AFAR D E.	
PA	(HUBE/) HUBERT R S.	
PA	(LEON/) LEONG K.	
PA	(RAIT/) RAITANO A B.	
PA	(SAFE/) SAFFRAN D C.	
XX		
PI	Afar DE, Hubert RS, leong K, Raitano AB, Saffran DC;	
XX		
DR	WPI; 2000-072832/06.	
DR	N-PSDB; AAZ49395, AAZ49396.	
XX		
PT	Novel proteins useful as diagnostic markers and therapeutic targets,	
PT	particularly for prostatic cancer.	
XX		
PS	Claim 1; Fig 1A; 83pp; English.	
XX		
CC	This sequence represents a novel human protein, STRAP-1 (serpentine	
CC	transmembrane antigen of the prostate). STRAP-1 is the prototype member	
CC	of the STRAP family of proteins (AAV58194-Y58197) which exhibit a high	
CC	degree of structural conservation, but which show no significant	
CC	structural homology to known human proteins. The STRAP-1 gene has been	
CC	localised to chromosome 7p22. STRAP-1 is thought to be a type IIIa	
CC	membrane protein and is expressed predominantly in prostate cells in	
CC	normal human tissues. Structurally, STRAP-1 is a 339 amino acid protein	
CC	characterised by six transmembrane domains and intracellular N- and C-	
CC	termini, suggesting that it folds in a "serpentine" manner into three	
CC	extracellular and two intracellular loops. STRAP-1 mRNA and protein	
CC	expression is maintained at high levels and throughout all stages of	
CC	prostate cancer. STRAP-1 mRNA and/or protein is also overexpressed in	
CC	certain other cancers, including bladder, colon, pancreatic and ovarian	
CC	cancer. The function of the STRAP proteins is not known. They may be ion	
CC	channels (from the presence of six transmembrane domains, a feature which	

CC	is shared by certain ion channels) or gap-junction proteins (from
CC	immunohistochemical staining). STRAP-1 and STRAP-2 are cell-surface
CC	tumour antigens. Immunisation with a STRAP protein induces cellular and
CC	humoral immune responses against STRAP-expressing cells. STRAP proteins
CC	may be used to identify specific-binding agents, to produce anticancer
CC	vaccines and to generate specific antibodies. The antibodies may be used
CC	for detection, prognosis, and monitoring of cancers (or susceptibility to
CC	cancer), as therapeutic inhibitors or to target therapeutic agents to
CC	their site of action. STRAP nucleic acids may be used for recombinant
CC	protein production, as diagnostic and prognostic reagents, for
CC	identifying STRAP-expressing cells for screening inhibitors of STRAP
CC	expression and for therapeutic modulation/inhibition of STRAP expression.
CC	Since high levels of STRAP proteins are exposed on the cell surface, they
CC	are easily targeted by systemically administered agents, and because
CC	they are expressed mainly on prostatic epithelial cells, agents targeted
CC	to them should have minimal side effects on other tissues
CC	
XX	
SQ	Sequence 339 AA:
Query Match	100.0%; Score 93; DB 3; Length 339;
Best Local Similarity	100.0%; Pred. No. 1.3e-05;
Matches 15; Conservative	0; Mismatches 0; Indels 0; Gaps 0
QY	1 WIDIKQFWYTPPTF 15
DB	277 WIDIKQFWYTPPTF 291
RESULT 6	
AAM01282	AAM01282 standard; protein; 339 AA.
XX	
AC	AAM01282;
XX	
DT	04-OCT-2001 (first entry)
DE	P789P amino acid sequence.
XX	
KW	Human; prostate cancer; prostate-specific; diagnosis; vaccine; cytostatic; gene therapy; metastasis.
OS	Homo sapiens.
XX	
PN	WO200151633-A2.
XX	
PD	19-JUL-2001.
XX	
PF	16-JAN-2001; 2001WO-US001574.
XX	
PR	14-JAN-2000; 2000US-00483672.
PA	(CORI-) CORIXA CORP.
PI	Xu J, Dillon DC, Mitcham JL, Harlocker SU, Jiang Y, Reed SG; Kalos MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skeiky YAW, Wang A, Meagher MJ;
DR	WPI; 2001-425873/45.
XX	
PT	New polynucleotide encoding a prostate-specific protein, for diagnosing, monitoring and treating prostate cancer in a patient and for use in vaccines.
PS	Claim 2; Page 510-512; 543pp; English.
XX	
CC	The present invention describes polynucleotide sequences (I) which encode
CC	prostate-specific proteins (II). (I) and (II) have cytostatic activity,
CC	and can be used in vaccine production and gene therapy. (I), (II), T cells
CC	antibodies to (II), fusion proteins comprising (II), and isolated T cells
CC	prepared using (I) or (II) are used treat cancer in a patient. (I) and
CC	the antibodies are also used in the detection of cancer in a patient. The
CC	cancer that is diagnosed or treated is particularly prostate cancer. (I)
CC	and (II) can be used in vaccines. The antibodies or (I) can be used for

CC monitoring the progression of cancer in a patient. (I) and (II) can also
CC be used to improve diagnostic and therapeutic methods for prostate
CC cancer. They can indicate the level of metastasis as well as the prostate
CC volume. AAH93357 to AAH93944 and AAM01115 to AAM0318 represent
CC polynucleotide and amino acid sequences used in the exemplification of
CC the present invention
XX

Sequence 339 AA:

Query Match 100.0%; Score 93; DB 4; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WIDIKQFWYTPPTF 15
Db 277 WIDIKQFWYTPPTF 291

RESULT 7

AAU69927 standard; protein; 339 AA.

AC AAU69927;
XX
DT 30-JAN-2002 (first entry)
XX
DE Human prostate cDNA encoded protein #72.
XX
KW Human; prostate cancer; cytostatic; immunostimulant; tumour; immunogen.
XX
OS Homo sapiens.
XX
PN WO200173032-A2.
XX
PD 04-OCT-2001.
XX
PF 27-MAR-2001; 2001WO-US009919.
XX
PR 27-MAR-2000; 2000US-00536857.
PR 09-MAY-2000; 2000US-00568100.
PR 12-MAY-2000; 2000US-00570737.
PR 13-JUN-2000; 2000US-00593793.
PR 27-JUN-2000; 2000US-00605783.
PR 09-AUG-2000; 2000US-00636215.
PR 29-AUG-2000; 2000US-00651236.
PR 06-SEP-2000; 2000US-00657279.
PR 02-OCT-2000; 2000US-00679426.
PR 10-OCT-2000; 2000US-00685166.
PR 09-NOV-2000; 2000US-00709729.
XX

(CORI-) CORIYA CORP.

XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
PI Li SX, Wang A, Skeiky YAM, Hepler WT, Henderson RA;
XX

WPI; 2001-639232/73.

DR N-PSDB; AAS64160.

PT New human prostate-specific polypeptides and polynucleotides useful for
XX the diagnosis and treatment of cancer, especially prostate cancer.

PS Claim 2; Page 549; 579pp; English.

XX The invention relates to isolated prostate-specific polynucleotides,
CC polypeptides, fusion proteins of the polypeptides, antibodies raised
CC against the polypeptides (or antigenic epitopes derived from them) and
CC antigen-presenting cells expressing the polypeptides. The antibodies are
CC useful for detecting the presence of cancer, especially prostate cancer.
CC The polypeptides, polynucleotides and the antigen-presenting cells are
CC useful for stimulating and/or expanding T cells specific for a tumour
CC protein, and for inhibiting the development of cancer especially prostate
CC cancer. Compositions comprising the polynucleotide and/or polypeptide are

CC useful for stimulating an immune response, and for treating cancer. The
CC oligonucleotide is useful for detecting cancer. The present sequence is a
CC prostate specific polypeptide of the invention
XX

Sequence 339 AA:

Query Match 100.0%; Score 93; DB 4; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WIDIKQFWYTPPTF 15
Db 277 WIDIKQFWYTPPTF 291

RESULT 8

AAM78845 standard; protein; 339 AA.

AC AAM78845;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 1507.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US004098.
XX
PR 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 01-SEP-2000; 2000US-0064936.
PR 15-SEP-2000; 2000US-00663561.
PR 20-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.
XX

(HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX

WPI; 2001-476283/51.

DR N-PSDB; AAK51978.

PT Nucleic acids encoding polypeptides with cytokine-like activities, useful
XX in diagnosis and gene therapy.

PS Claim 20; Page 3800-3801; 6221pp; English.

XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activity/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAM8020) are omitted as the relevant pages from the

```
CC sequence listing were missing at the time of publication
XX
SQ Sequence 339 AA;
Query Match 100.0%; Score 93; DB 4; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 WIDIKQFWYTPPTF 15
DQ 277 WIDIKQFWYTPPTF 291
RESULT 9
ABU71818
XX ABU71818 standard; protein; 339 AA.
XX
AC ABU71818;
XX
DT 10-JUN-2003 (first entry)
XX
DE Prostate cancer associated protein #53.
XX
KW Prostate cancer; vaccine; gene therapy; cytostatic; fusion protein;
KW immunogen; cancer; prostate specific antigen; PSA;
KW prostatic acid phosphatase; PAP; prostate specific membrane antigen;
KW PSMA.
XX
OS Homo sapiens.
XX
PN US2002192763-A1.
XX
PD 19-DEC-2002.
XX
PF 29-JUN-2001; 2001US-00895793.
XX
PR 04-OCT-1999; 99US-0157455P.
PR 04-OCT-2000; 2000US-00679272.
PR 28-MAR-2001; 2001US-00822827.
XX
XX (XUUU/) XU J.
XX PA (DILL/) DILLON D C.
XX PA (MITC/) MITCHAM J L.
XX PA (HARL/) HARLOCKER S L.
XX PA (JIANG/) JIANG Y.
XX PA (KALO/) KALOS M D.
XX PA (FRANG/) FRANGER G R.
XX PA (REIT/) REITTER M W.
XX PA (STOL/) STOLK J A.
XX PA (DAYC/) DAY C H.
XX PA (VEDV/) VEDVICK T S.
XX PA (CART/) CARTER D.
XX PA (LISX/) LI S X.
XX PA (WANG/) WANG A.
XX PA (SKEI/) SKEIKY Y A W.
XX PA (HEPL/) HEPLER W T.
XX PA (HEND/) HENDERSON R A.
XX PA (HURA/) HURAL J.
XX PA (MCNE/) MCNETIL P D.
XX PA (HOUN/) Houghton R L.
XX PA (DBAS/) Y DE BASOLS C V.
XX PA (FOYT/) FOY T M.
XX
XX Xu J, Dillon DC, Mitcham JL, Harlocker SH, Jiang Y, Kalos MD;
XX Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
XX Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA, Hural J;
XX McNeill PD, Houghton RL, Y De BasolsCV, Foy TM;
XX
XX WPI; 2001-245062/25.
XX
XX N-PSDB; ACAA59965.
XX
XX Prostate specific protein and its encoding polynucleotide, useful for the
XX treatment and diagnosis of prostate cancer.
XX
PT
```

```
XX
XX Example 5; SEQ ID NO 879; 85pp; English.
XX
XX The invention describes a fusion protein comprising at least one amino
XX acid sequence of immunogenic portions of any of the 3 sequences not
XX defined in the specification, or sequences having at least 70 or 90 %
XX sequence identity to any one of the 35 sequences defined in the USPTO web
XX site, which is encoded by any of the 4 nucleotide sequences not defined
XX in the specification. The fusion protein, composition and methods are
XX useful for diagnosing, preventing and/or treating cancer, particularly
XX prostate cancer. The proteins are useful as markers to indicate the
XX presence or absence of cancer. This is the amino acid sequence of a
XX prostate cancer therapy associated protein. Note: The sequence data for
XX this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from the US patent office at
XX seqdata.uspto.gov/sequence.html?DocID=US20020192763
XX
XX Sequence 339 AA;
SQ
Query Match 100.0%; Score 93; DB 4; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 WIDIKQFWYTPPTF 15
DQ 277 WIDIKQFWYTPPTF 291
RESULT 10
ABB95387
XX ABB95387 standard; protein; 339 AA.
XX
AC ABB95387;
XX
DT 19-JUL-2002 (first entry)
XX
DE Human P789P protein SEQ ID NO 879.
XX
KW Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;
KW gene therapy.
XX
XX Homo sapiens.
XX
XX US2002022248-A1.
XX
XX 21-FEB-2002.
XX
XX 12-JAN-2001; 2001US-00759143.
XX
XX 25-FEB-1997; 97US-00806099.
XX 01-AUG-1997; 97US-00904804.
XX 10-FEB-1998; 98US-00020956.
XX 25-FEB-1998; 98US-00030607.
XX 14-JUL-1998; 98US-00115453.
XX 23-SEP-1998; 98US-00159812.
XX 15-JAN-1999; 99US-00232149.
XX 09-APR-1999; 99US-00288946.
XX 13-JUL-1999; 99US-00352616.
XX 12-NOV-1999; 99US-00439313.
XX 18-NOV-1999; 99US-00443586.
XX 14-JAN-2000; 2000US-00483672.
XX 27-MAR-2000; 2000US-00536857.
XX 09-MAY-2000; 2000US-00568100.
XX 12-MAY-2000; 2000US-00570737.
XX 13-JUN-2000; 2000US-00593793.
XX 27-JUN-2000; 2000US-00605783.
XX 10-AUG-2000; 2000US-00636215.
XX 29-AUG-2000; 2000US-00651236.
XX 06-SEP-2000; 2000US-00657279.
XX 02-OCT-2000; 2000US-00679426.
XX 10-OCT-2000; 2000US-00685166.
XX
XX (XUUU/) XU J.
XX
```

```

PR 30-APR-2001; 2001US-00847046.
PR 04-MAY-2001; 2001US-0285859P.
PA (EOSB-) EOS BIOTECHNOLOGY INC.
PI Gish KC, Mack DH, Wilson KE, Afari D, Hevezi P;
PX WPI; 2002-471335/50.
DR N-PSDB; ABK92128.
XX
PT Detecting a prostate cancer-associated transcript in a cell in a patient,
PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,
PT by determining if prostate cancer-associated genes are expressed in a
PT prostate tissue.
XX
PS Claim 27; Page 312; 436pp; English.
XX
CC The present invention relates to methods of detecting a prostate cancer-
CC associated transcript in a cell from a patient. The method comprises
CC contacting a biological sample from the patient with prostate cancer-
CC associated polynucleotides (designated PC genes) that selectively
CC hybridize to a sequence that is at least 80% identical to them. The
CC prostate cancer-associated polynucleotide sequences are differentially
CC expressed in prostate tumour tissue or in prostate cancer and are derived
CC from the tissues of various organisms such as humans or other mammals
CC (e.g. mice, sheep and dogs). The methods of the invention are useful for
CC diagnosing and treating prostate cancer in mammals. The prostate cancer-
CC associated genes are useful for diagnosing or treating prostate cancer,
CC as well as for identifying modulators of prostate cancer or agents that
CC inhibit prostate cancer. The nucleic acid sequences are particularly
CC useful in gene therapy, as a vaccine or in antisense applications.
CC ABG61800-ABG61944 represent prostate cancer-associated proteins.
SQ
SQ Sequence 339 AA;
SQ
Query Match 100.0%; Score 93; DB 5; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 WIDIKQFWYTPPTP 15
Db 277 WIDIKQFWYTPPTF 291
RESULT 12
ABU98383
ID ABU98383 standard; protein, 339 AA.
XX AC
XX ABU98383;
XX
DT 31-JUN-2003 (first entry)
XX
DE Novel human gene STEAP-1 variant 1.
XX
KW STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
KW cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
KW vaccine.
XX
XX Homo sapiens.
XX OS
XX MO2003022995-A2.
XX PN
XX PD 20-MAR-2003.
XX PF 06-SEP-2002; 2002MO-US028371.
XX PE 06-SEP-2001; 2001US-0317840P.
XX PR 05-APR-2002; 2002US-0370387P.
XX
XX PA (AGEN-) AGENSYS INC.
XX
XX Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;
XX

```

DR WPI; 2003-313240/30.
DR N-PSDB; ACD02597.
XX
PT New composition comprising a substance that modulates the status of a
PT STEAP-1-related protein, useful for treating and detecting cancer.
XX
PS Example 2; Fig 2A; 248pp; English.
XX
CC The invention describes a composition comprising a substance that
CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
CC any of the 15 sequences of 259 amino acids, given in the specification,
CC or a molecule that is modulated by the protein, where the status of the
CC cell that expresses the protein is modulated. The compositions, proteins,
CC polynucleotides and methods are useful for treating and detecting cancer.
CC The STEAP-1-related proteins are useful for generating cancer vaccines.
CC The polynucleotides are useful as tools for delineating, with greater
CC precision, cytogenetic abnormalities in the chromosomal region that
CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
CC the amino acid sequence of a variant of human six transmembrane
CC epithelial antigen of the prostate or STEAP-1
XX
SQ Sequence 339 AA;
XX
Query Match 100.0%; Score 93; DB 6; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 WIDIKQFWYTPPTF 15
|||
277 WIDIKQFWYTPPTF 291
XX
DB
RESULT 13
ID ABU98414 standard; protein; 339 AA.
AC ABU98414;
XX
XX 31-JUL-2003 (first entry)
DT
XX
DE STEAP-1 variant 1.
XX
XX STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
XX cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
XX vaccine.
XX
OS Homo sapiens.
XX
XX W02003022995-A2.
XX
XX 20-MAR-2003.
PD
XX
PF 06-SEP-2002; 2002WO-US028371.
XX
XX 06-SEP-2001; 2001US-0317840P.
PR 05-APR-2002; 2002US-0370387P.
XX
XX (AGEN-) AGENSYS INC.
PA
XX Farris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;
PI WPI; 2003-313240/30.
XX
XX WPI; 2003-313240/30.
DR
XX
PT New composition comprising a substance that modulates the status of a
PT STEAP-1-related protein, useful for treating and detecting cancer.
XX
PS Example 53; Page 173; 248pp; English.
XX
XX The invention describes a composition comprising a substance that
XX modulates the status of a protein (I) of 340 or 283 amino acids, or of
XX any of the 15 sequences of 259 amino acids, given in the specification,
XX or a molecule that is modulated by the protein, where the status of the
XX cell that expresses the protein is modulated. The compositions, proteins,
XX

CC polynucleotides and methods are useful for treating and detecting cancer.
CC The STEAP-1-related proteins are useful for generating cancer vaccines.
CC The polynucleotides are useful as tools for delineating, with greater
CC precision, cytogenetic abnormalities in the chromosomal region that
CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
CC the amino acid sequence of a variant of human six transmembrane
CC epithelial antigen of the prostate or STEAP-1
XX
SQ Sequence 339 AA;
XX
Query Match 100.0%; Score 93; DB 6; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 WIDIKQFWYTPPTF 15
|||
277 WIDIKQFWYTPPTF 291
XX
DB
RESULT 14
ID ABU98430 standard; protein; 339 AA.
AC ABU98430;
XX
XX 31-JUL-2003 (first entry)
DT
XX
DE STEAP-1 variant 8P1D4 v.1 #1.
XX
XX STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
XX cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
XX vaccine.
XX
OS Homo sapiens.
XX
XX W02003022995-A2.
XX
XX 20-MAR-2003.
PD
XX
PF 06-SEP-2002; 2002WO-US028371.
XX
XX 06-SEP-2001; 2001US-0317840P.
PR 05-APR-2002; 2002US-0370387P.
XX
XX (AGEN-) AGENSYS INC.
PA
XX Farris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;
PI WPI; 2003-313240/30.
XX
XX WPI; 2003-313240/30.
DR
XX
PT New composition comprising a substance that modulates the status of a
PT STEAP-1-related protein, useful for treating and detecting cancer.
XX
PS Example 9; Fig 13A; 248pp; English.
XX
XX The invention describes a composition comprising a substance that
XX modulates the status of a protein (I) of 340 or 283 amino acids, or of
XX any of the 15 sequences of 259 amino acids, given in the specification,
XX or a molecule that is modulated by the protein, where the status of the
XX cell that expresses the protein is modulated. The compositions, proteins,
XX polynucleotides and methods are useful for treating and detecting cancer.
XX The STEAP-1-related proteins are useful for generating cancer vaccines.
XX The polynucleotides are useful as tools for delineating, with greater
XX precision, cytogenetic abnormalities in the chromosomal region that
XX encodes STEAP-1 that may contribute to the malignant phenotype. This is
XX the amino acid sequence of a variant of human six transmembrane
XX epithelial antigen of the prostate or STEAP-1
XX
SQ Sequence 339 AA;
XX
Query Match 100.0%; Score 93; DB 6; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WIDIKQFVWYTPPTF 15
 |||||
 Db 277 WIDIKQFVWYTPPTF 291

RESULT 15

ABR54499

ID ABR54499 standard; protein; 339 AA.

AC ABR54499;

DT 28-AUG-2003 (first entry)

DE Prostate tumour specific protein SEQ ID 879.

KW Cytostatic; gene therapy; prostate-specific protein; PSP; human;
 immune response; prostate cancer.

OS Homo sapiens.

PN W0200289747-A2.

PD 14-NOV-2002.

PF 09-MAY-2002; 2002WO-US014753.

PR 09-MAY-2001; 2001US-00852911.

PR 29-JUN-2001; 2001US-00895814.

PR 10-DEC-2001; 2001US-00012896.

(CORI-) CORIXA CORP.

XU J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Henderson RA,
 Kalos MD, Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS,
 Carter D, Li SX, Wang A, Skeiky YAW, Hepler WT, Hurai J,
 McNeill PD, Houghton RU, Vinals Y De BassolsC, Foy TM, Watanabe Y,
 Deng T;

WPI; 2003-167130/16.

PT New prostate-specific proteins and genes, useful in gene therapy,
 particularly for stimulating an immune response in a patient, or treating
 prostate cancer in a patient, as well as for diagnosing prostate cancer
 in a patient.

Example 5; Page 618-619; 691pp; English.

CC The present invention relates to novel prostate-specific proteins (PSP)
 and their coding sequences. The PSPs and their coding sequences are
 useful for stimulating an immune response in a patient, or for treating
 prostate cancer in a patient and for determining, detecting or diagnosing
 the presence of a cancer in a patient. The present sequence was used to
 illustrate the invention

SQ Sequence 339 AA;

Query Match

100.0%; Score 93; DB 6; Length 339;

Best Local Similarity 100.0%; Pred. NO. 1.3e-05;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WIDIKQFVWYTPPTF 15
 |||||
 Db 277 WIDIKQFVWYTPPTF 291

Search completed: July 12, 2004, 14:07:24
 Job time : 33.9651 secs

This Page Blank (uspto)

This Page Blank (uspto)